The Rational Hypocrite: Informal Argumentation and Moral Hypocrisy

Tage S. Rai (tage.rai@gmail.com)
Department of Psychology
University of California, Los Angeles
Los Angeles, CA 90095 USA

Keith J. Holyoak (holyoak@lifesci.ucla.edu)
Department of Psychology
University of California, Los Angeles
Los Angeles, CA 90095 USA

Abstract

We suggest that in some instances the apparent logical inconsistency of moral hypocrisy stems from different evaluations of a weak argument, rather than dishonesty per se. Extending Corner, Hahn, and Oaksford's (2006) analysis of slippery slope arguments, we propose that inferences of hypocrisy depend on perceived similarity of actions to previous standards. In Experiment 1, dissimilar actions were rated as less hypocritical than their similar counterparts. If observers are choosing between competing theories (i.e., hypocrisy or legitimate dissimilarity), evidence of self-serving motives will positively support inferences of hypocrisy independent of changes in similarity. In Experiment 2, we manipulated potential self-serving interests that an action would produce while keeping similarity between cases identical. Actions that would result in a beneficial outcome for the actor were seen as more hypocritical than their non-self-serving counterparts. These results support the possibility that Bayesian analyses of weak arguments have implications for assessing moral reasoning.

Keywords: moral reasoning; hypocrisy; Bayesian inference; argumentation

Introduction

"I respect the jury's verdict. But I have concluded that the prison sentence given to Mr. Libby is excessive." With these words former President George W. Bush commuted the sentence of I. Lewis "Scooter" Libby, Jr., for obstruction of justice and leaking the identity of CIA operative Valerie Plame. Critics of the decision noted that Libby actually had received the minimum sentence allowable for his offense under the law, and that many of Libby's supporters, including the Bush administration, were actively pressing for mandatory minimum sentencing laws at a national level (Goodman, 2007). Accordingly, critics of the decision saw it as a textbook case of moral hypocrisy: different rules were being applied to Bush's underling, Libby, than to everyone else in the United States.

As the example illustrates, moral hypocrisy is typically viewed as an ethical accusation: someone is applying different moral standards to essentially identical cases, dishonestly claiming that one action is acceptable whereas otherwise equivalent actions are not (Valdesolo & DeSteno, 2007). In this framework, accusations of hypocrisy entail an inherent *logical inconsistency* with *moral implications*, in that the morally hypocritical action is of the same category as previous standards, and therefore should not be treated any differently according to some norm of justice. The assumption (at least implicitly) is that the hypocrite is being dishonest, or at least self-deceptive. The hypocrite must be aware (or should be aware) of the logical inconsistency, and is therefore committing a falsehood (Batson et al., 1997; for a review of hypothesized bases for moral disagreements, see Rai & Fiske, 2011).

Although people no doubt sometimes engage in such genuine hypocrisy, there may be an alternative explanation of *apparent* hypocrisy, one that leaves open the possibility that the (perceived) hypocrite is both honest and rational. This possibility arises as an extension of recent Bayesian analyses of informal argumentation. Hahn and Oaksford (2007) have argued that many seeming fallacies of informal reasoning (e.g., arguments from ignorance, circular arguments, and slippery slope arguments), are not in fact illogical in all cases, but rather are more or less rationally persuasive depending on the content of the argument. For example, Corner, Hahn and Oaksford (2006) analyzed the slippery slope argument, in which a proposed action A is criticized because it may lead to an undesirable outcome B. Corner et al. (2006) argued that such arguments hinge in part on the conditional probability that undesirable outcome B will in fact occur if action A is taken. The conditional probability value,[1] depends in part on the prior probability that action A will expand a category boundary far enough to include B in the future. In general, the greater the perceived similarity of A and B, the greater the probability that including A will expand the category boundaries to incorporate B.

For example, the claim that legalizing marijuana use might lead to the eventual legalization of heroin may seem less persuasive than the claim that legalizing cocaine might lead to the legalization of heroin, because heroin seems more similar to cocaine than to marijuana. In general, to the extent the slope in fact seems "slippery" (high similarity of

---

[1] More precisely, the relevant quantity would presumably be a function of the difference between the probability of B conditional on A versus not-A (intuitively, a judgment of the probability that A will causally lead to B).

A to B), a slippery slope argument against taking action A may be rationally persuasive.

The logical structure of an allegation of moral hypocrisy is related to the structure of slippery slope arguments. A critic of a slippery slope argument is likely to claim that cases A and B do not (and will not) belong to the same category. In contrast, one who alleges moral hypocrisy is likely to claim that A and B in fact belong to the same category, and hence should be viewed as the "same" from a moral perspective. Thus when we perceive someone as a hypocrite, it is not necessarily because there is an inherent logical inconsistency in their argument, and their claim may not be motivated by their selfish desire to do A while blocking B. An alternative possibility is that we estimate the similarity between A and B to be high, whereas the (alleged) hypocrite's premise is that they are actually dissimilar, so that doing A does not morally require acceding to B.

The close link between slippery slope arguments and allegations of moral hypocrisy can be illustrated by returning to the Scooter Libby commutation. We could turn the critique of Bush's action into a slippery slope argument by claiming that commuting Libby's sentence would morally require us to commute everyone else convicted of similar crimes because Libby's crime was no less heinous than theirs (i.e., the slope is so slippery as to be impossible to resist). The stronger this slippery slope argument seems, the weaker is Bush's perceived defense against the allegation of moral hypocrisy. However, perhaps Bush in fact viewed Libby's offense as dissimilar to the crimes of others convicted of the same charges (because, for example, Libby was a "patriot" motivated by "noble" motives, such as increasing the security of the United States by advancing the agenda of the Bush administration in overcoming opposition to its plan to invade Iraq). Given such prior beliefs, Bush could rationally decide that commuting Libby (based on his exculpatory circumstances and exemplary character) was rationally and morally consistent with not commuting the sentences of others convicted of similar crimes.

To take a different example, a critic might label a pro-choice individual as hypocritical if that person thinks it should be illegal for pregnant women to drink alcohol. From this critic's point of view, both issues may be seen as involving a woman's right to control her own body; hence if a woman has the right to choose abortion, she should surely have the right to drink alcohol. However, the allegation of hypocrisy could be countered by arguing that the actions are in fact quite dissimilar in relevant respects (e.g., demanding that a woman maintain a pregnancy may have much greater negative utility for her than demanding that she stop drinking for the same period of time). As this example, illustrates, making a distinction between the utilities of A and B is an important way to show that A and B are dissimilar in a morally relevant way.

That apparent moral hypocrisy may sometimes be the result of disagreement over content does not imply that there is no "genuine" moral hypocrisy stemming from deception and differential treatment in the service of selfish ends. For example, critics of Bush's decision to commute Libby's sentence have noted that if Bush had given Libby a full pardon, Libby would have had to relinquish his Fifth Amendment rights, whereas commuting the sentence maintained those rights (Goodman, 2007). Thus, although Bush claimed that he chose commutation rather than a full pardon in order to appropriately punish Libby, in fact it was in Bush's own best interests to commute Libby rather than to either pardon him or allow him to go to jail, as either alternative would have increased the probability of Libby cooperating with prosecutors and revealing the possible involvement of President Bush or Vice-President Cheney in Libby's crimes.

Such self-serving interests can be modeled in Bayesian terms when assessing the probability of moral hypocrisy. For the critical observer faced with Bush's decision, the two competing hypotheses are that Bush had different prior beliefs that lead to his decision, or that Bush believed that criminals should go to jail and that Libby was a criminal, but hypocritically commuted his sentence. Viewed in a Bayesian framework, evidence that Bush may have had self-serving reasons to commute Libby's sentence would count as positive evidence in support of the hypothesis that Bush was a hypocrite, independent of any changes in the similarity between Libby's case and that of other convicts.

In the present paper we test our proposal that perceptions of moral hypocrisy depend on underlying probabilities related to content. In Experiment 1 we manipulated the similarity between a precedent action A and a proposed action B. We hypothesized that as A and B became more similar, participants would be increasingly likely to view treating them differently as morally hypocritical. In Experiment 2 we manipulated the self-serving benefits to the actor while keeping the similarity identical between precedent action A and proposed action B. We hypothesized that as the potential for self-serving interests increased, participants would be increasingly likely to view treating the actions differently as morally hypocritical.

Experiment 1

Participants in both experiments were recruited via the Internet, and completion of the questionnaires was voluntary. Questionnaires were posted on the Craigslist website under their "volunteers" section in Los Angeles, New York, and Chicago. After giving consent to participate, each participant was randomly assigned to a condition. The IP addresses of participants' computers were recorded to ensure that participants did not participate in the study multiple times.

Experiment 1 employed a within-subjects design in which participants (n = 59) were presented with two different vignettes, each representing a different experimental condition, counterbalanced for order. Each vignette was varied to experimentally manipulate the similarity between a precedent action A and a proposed

action B. Similarity varied along two dimensions: surface similarity based on features and utility based on consequences. A given participant read one vignette in which the two actions were made to appear quite similar on the surface and the utility of the outcome was relatively more negative (from the participants' perspective), and another vignette in which the two actions were made to appear less similar and the utility of the outcome less negative. Participants were asked to rate how morally hypocritical it would be to adopt the proposed action B on a Likert scale ranging from 1 (not hypocritical at all) to 7 (completely hypocritical). The information in parentheses refer to the low similarity condition:

For several years John Smith ran a successful investment company, but during the recession he made some bad investments. Realizing that all of his savings were going to disappear, John made a desperate move to embezzle some of the money from his company. He knew it was wrong, but gave in to temptation. John was caught, and in his state his crime carried a mandatory sentence of 5 years in prison. John's lawyer has asked the judge to reduce the sentence to 2 years. Other convicted felons are in prison for monetary crimes, such as theft and money laundering *(violent crimes, such as homicide and assault).*

How morally hypocritical would it be for the judge to reduce John's sentence to two years when other convicts are sentenced to more time?

In Norway, they have passed a law making it illegal to drink alcohol while pregnant because of the side-effects that it may have on the developing infant. When pregnant women drink alcohol, the alcohol is passed directly to the infant, and can lead to severe side-effects, including fetal-alcohol syndrome. Some people are advocating that this class of laws be expanded to include caffeine consumption during pregnancy. Caffeine consumption is passed directly to the infant. It can lead to severe side-effects, including type 2 diabetes. *(Caffeine consumption is passed indirectly to the infant. It can lead to some side-effects, including a slightly higher likelihood of consuming caffeine in the future.)*

How morally hypocritical is it for the law to permit caffeine consumption while restricting alcohol consumption during pregnancy?

According to our hypothesis, reducing John's sentence to 2 years should be seen as less hypocritical when his crime is markedly different from that of other criminals in a way that reduces the negative utility of John's crime to society (monetary rather than violent). Similarly, it should be seen as less hypocritical to allow pregnant mothers to consume caffeine but not alcohol when the effects of caffeine are different from those of alcohol (indirect rather than direct causal pathway) and less harmful to the infant (less negative utility).

Results

ANOVA analyses revealed that ratings of hypocrisy were significantly higher when the precedent action A and the proposed action B were highly similar based on surface and utility similarity ($M = 4.54$) than when the actions were more dissimilar ($M = 3.27$), $F (1, 55) = 7.69$, $p < .01$ (see Figure 1). Ratings did not differ across the two vignettes, $F (1, 55) < 1$. More importantly, there was no interaction between vignette and experimental manipulation, $F (1, 55) = 1.85$, $p = .18$, indicating that the basic effect was observed for both vignettes (see Figure 1). There was also no effect of presentation order, $F (1, 55) = 1.43$, p = .24.
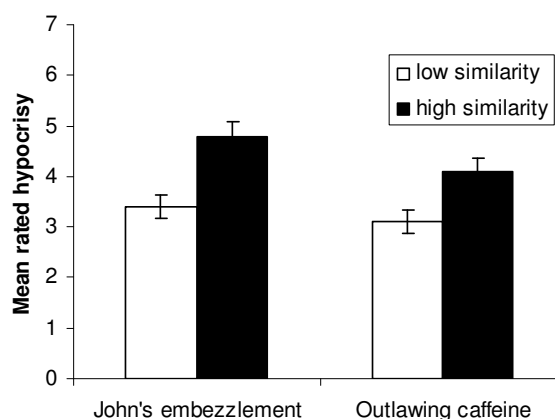

Figure 1: Mean ratings of moral hypocrisy as a function of similarity between cases, for each of two vignettes (Experiment 1). A rating of one corresponded to "not hypocritical at all" while a rating of seven corresponded to "completely hypocritical". Error bars indicate one standard error of the mean.

Experiment 2

Whereas Experiment 1 demonstrated the effects of varying similarity on accusations of hypocrisy, in Experiment 2 we varied the likelihood of self-serving motives while keeping similarity between cases constant.

Experiment 2 employed a between-subjects design in which participants were presented with one vignette that was varied to experimentally manipulate the benefit an actor would likely receive from performing an action. Benefit was manipulated via a financial incentive for performing the action. Specifically, a given participant read a vignette in which a governor can choose to reduce a man's sentence. Participants were either told that the man comes from a rich family (n = 22) or a poor family (n = 24):

James has been convicted of vehicular homicide due to killing a homeless person while driving drunk. There were extenuating circumstances, as James had no prior record, visibility was particularly poor due to inclement weather, and it was not a direct collision. The victim was not crossing at a cross-walk and was in poor health. However, because of mandatory sentencing laws in the state, the judge had to sentence James to seven years in prison. James has submitted a request to have his sentence reduced to three years. However, the governor of the state has been a strong proponent of minimum sentencing laws. James comes (does not come) from a very rich family, and they would (not) likely become large donors to the governor if he pardoned their son.

How morally hypocritical would it be for the governor to reduce the sentence James received for his crime?

According to our hypothesis, reducing James' sentence to 3 years should be seen as less hypocritical when his family is poor, because observers would perceive that the governor has no self-serving interest in taking the action

Results

Accusations of hypocrisy were significantly higher among participants who were told James came from a rich family (m = 5.64) than among participants who were told James came from a poor family (m = 3.58), $t(44) = 3.76$, $p < .001$ (see Figure 2). Thus, information about the potential for selfish gain supported accusations of hypocrisy independent of any changes to the similarity between James' crime and the crimes of others.
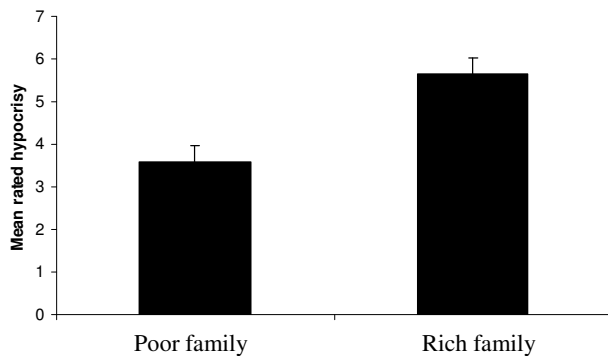

Figure 2: Mean ratings of moral hypocrisy as a function of potential for selfish gain between cases (Experiment 2). A rating of one corresponded to "not hypocritical at all" while a rating of seven corresponded to "completely hypocritical". Error bars indicate one standard error of the mean.

Discussion

The present findings demonstrate that apparent moral hypocrisy varies as a function of underlying probabilities related to content. Specifically, observers have two competing hypotheses for analyzing an actor's actions: Either he is a dishonest hypocrite or the proposed action B is actually different from precedent action A. As precedent action A and proposed action B become more similar, perceptions of moral hypocrisy will increase (Experiment 1) because the likelihood of the two actions being in the same category is increased. Perceptions of moral hypocrisy will also increase independent of changes in similarity if the potential for selfish gain is believed to be present (Experiment 2), as this will count as positive evidence in support of the hypothesis that the actor is a hypocrite.

These results extend work on other forms of informal argumentation (Hahn & Oaksford, 2007), and represent the first attempt to analyze moral hypocrisy in Bayesian terms.

Rather than being based on illogical motivated cognition, at least some apparent instances of moral hypocrisy may be rooted in Bayesian notions of argument strength. Differences in prior beliefs can lead to conflicting interpretations of argument strength and consequent accusations of hypocrisy and attempts at rebuttal. By linking moral hypocrisy to Bayesian decision theory, we extend recent work that has argued for more domain-general accounts of moral cognition (Rai & Holyoak, 2010; Waldmann & Dieterich, 2007; Waldmann, Nagel, & Wiegmann, in press).

An interesting question is whether people are sensitive to the roles of similarity and motives in inferences toward hypocrisy. If so, one might predict that genuine hypocrites will adjust their beliefs regarding the similarity between cases and the potential gains they will receive from taking the proposed action. Future studies should investigate how estimates of underlying probabilities related to similarity and gain change under conditions that are likely to lead to hypocrisy, such as when making judgments of others actions versus one's own actions (Valdesolo & Desteno, 2008).

References

Batson, D., Kobrynowicz, D., Dinnerstein, J., Kampf, H., & Wilson, A. (1997). In a very different voice: Unmasking moral hypocrisy. *Journal of Personality and Social Psychology*, *72*, 1335-1348.

Corner, A., Hahn, U., & Oaksford, M. (2006). The slippery slope argument—Probability, utility & category reappraisal. In R. Sun (Ed.), *Proceedings of the 28th Annual Meeting of the Cognitive Science Society* (pp. 1145–1150). Mahwah, NJ: Erlbaum.

Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review, 114,* 704-732.

Goodman, A. (2007). Commuting sentence, Bush spares Libby from 30 month jail-term. *Democracy Now: The War and Peace Report.* Retrieved from http://www.democracynow.org/2007/7/3/commuting_sentence_bush_spares_libby_from.

Rai, T. S. & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review, 118,* 57-75.

Rai, T. S., & Holyoak, K. J. (2010). Moral principles or consumer preferences? Alternative framings of the trolley problem. *Cognitive Science*, *34*, 311-321.

Valdesolo, P., & DeSteno, D. (2007). Moral hypocrisy: Social groups and the flexibility of virtue. *Psychological Science, 18, 689-690.*

Valdesolo, P., & Desteno, D. (2008). The duality of virtue: Deconstructing the moral hypocrite. *Journal of Experimental Social Psychology, 44,* 1334-1338.

Waldmann, M. R., & Dieterich, J. H. (2007). Throwing a bomb on a person versus throwing a person on a bomb: Intervention myopia in moral intuitions. *Psychological Science, 18,* 247-253.

Waldmann, M. R., Nagel, J., & Wiegmann, A. (in press).Moral judgment. In K.J. Holyoak & R. G. Morrison (Eds.),*The Oxford handbook of thinking and reasoning.* New York: Oxford University Press.