Predicting Moral Judgments of Corporate Responsibility with Formal Decision Heuristics

Anna Coenen (coenen@mpib-berlin.mpg.de)
Max Planck Institute for Human Development, Center for Adaptive Behavior and Cognition
Lentzeallee 94, 14195 Berlin, Germany

Julian N. Marewski (marewski@mpib-berlin.mpg.de)
Max Planck Institute for Human Development, Center for Adaptive Behavior and Cognition
Lentzeallee 94, 14195 Berlin, Germany

Abstract

We address the questions (1) how moral decision-making can be formally modeled using established decision models and (2) which of these models are the most accurate in predicting moral judgments. We conducted an experiment with a comparison task in which people had to decide which of two companies behaved more fairly in dealing with the consequences of the current financial crisis. We modeled these judgments of fairness (i) using a compensatory weighted additive model (WADD), (ii) a unit weight linear heuristic (UWL), and (iii) a noncompensatory decision rule (LEX). All strategies could predict people's actual decisions above chance level for a majority of the subjects. This lends support to the modeling approach to moral decision-making.

Keywords: moral heuristics; simple heuristics; moral psychology; take-the-best.

Introduction

In the late summer of the year 2008, financial markets all over the world dropped to historic lows which drove many Western societies into the most severe financial crisis since the Great Depression. The media, as well as public and political discussion remained looking for the scapegoat; eyes turned on CEOs and bankers all over the world in trying to find an explanation of the rapidly evolving recession. Major companies were faced with the difficult task of justifying sudden bankruptcy, massive rationalization programs, and job cuts.

Preserving trust in public opinion, customers, and stakeholders became a critical point in dealing with the difficult circumstances of plunging stock markets and decreasing consumption. The debate about corporate responsibilities and social justice that evolved highlights the moral significance that the crisis management of major companies has for a greater part of the public.

Evaluating the integrity and responsibility of companies, banks, government, or other public institution plays a great part in our every day dealing with morality. The current study seeks to model the moral evaluation processes that underlie normative judgments such as those involved in assessing corporate crisis management.

With respect to current theories of moral decision-making, our study has three aims: (1) To use formal models for describing and predicting moral decisions, (2) to spell out the concept of moral intuitions in terms of the *fast and frugal heuristics* approach to judgment and decision-making (e.g., Gigerenzer, Todd, and the ABC Research Group, 1999), and (3) to focus on realistic normative decision problems.

First, recent work in moral psychology has stressed the importance of *intuition* for moral decision-making (Greene & Haidt, 2002; Haidt, 2001; Hauser, Cushman, Young, Jin, & Mikhail, 2007). According to this work, people often make moral judgments quickly and unconsciously, without engaging in explicit deliberative reasoning. Only after a judgment has already been reached do they construct reasons for their already-made decisions. If moral decisions have little to do with our post-hoc rationalizations of them, then this leads to the question what the processes and reasons are that underlie moral judgments. Formal models of moral decision-making can help to explicate the notion of moral intuition and draw attention to the reasoning processes involved in moral judgment (Gigerenzer, 2008). Moreover, by building formal models of people's judgment processes the precision and predictive power of corresponding theories can be increased, which in turn can foster cumulative theory building and, in doing so, aid scientific progress (for an overview of some of the advantages of formal modeling, see Marewski & Olsson, 2009). To our best knowledge, up to today there are not many attempts to study moral judgments in terms of formal, cognitive decision models. In this study, we therefore aim at testing the ability of different models to predict moral decisions.

Second, one way of spelling out the fast and intuitive nature of moral judgments is to understand them as being produced by simple rules of thumb, or *heuristics* (Gigerenzer, 2007, 2008; Sunstein, 2005). The fast and frugal heuristics research program has shown that such heuristics can successfully predict people's decisions when they answer questions about *matters of fact*, for instance, when inferring which of two cities is larger, or when making estimates about some other objective criterion (e.g., Bergert & Nosofsky, 2007; Bröder & Gaissmaier, 2007; von Helversen & Rieskamp, 2008; Mata, Schooler, & Rieskamp, 2007; Rieskamp & Hoffrage, 2008; Rieskamp & Otto, 2006; for a recent overview see e.g., Marewski, Galesic, Gigerenzer, 2009). Here, we study if and how some well-known examples of these heuristics and related earlier

models (see e.g., Dawes, 1979; Dawes & Corrigan, 1974; Payne, Bettman & Johnson, 1993) can predict *normative* decisions. The models include a *lexicographic decision rule* (*LEX*), a *unit weight linear heuristic* (*UWL*), as well as a more complex *weighted additive model* (*WADD*).

Third, in studying moral decision making, we focus on real-world decision problems rather than constructed moral dilemmas. Frequently, research in experimental moral psychology makes use of artificial situations, such as *trolley problems*, to study moral decision-making (Greene & Haidt, 2002; Greene, Sommerville, Nytrom, Darley, & Cohen, 2001; Mikhail, 2007). These dilemmas portray situations of a kind that many people are most likely to be unfamiliar with. For example, presumably very few people have ever been confronted with a situation where they had to decide whether to let four people die who are lying on a train track, or whether to redirect an oncoming train to save them and, in doing so, kill one other person. (This is a standard trolley dilemma.) However, moral judgment needs to be understood as an ability that is to a great extend shaped by the natural (i.e., social) environment of a person; only if the decision problems used for studying moral judgments are familiar to the participants, people can rely on their every day moral intuitions. In our study, we therefore chose examples of problems that people also encounter in the real world, such as the evaluation of corporate justice and responsibility.

In what follows, we will first present the three models that we tested. In an experiment, we will then evaluate each model's performance in predicting participants' moral judgments about the fairness of companies' behavior.

Model Introduction

All models were tested using a *two-alternative forced-choice task* in which subjects had to decide which of two companies makes the fairer decisions. For instance, one task was to decide whether company *A* or company *B* behaved more fairly in dealing with the consequences of the financial crisis: To compensate for their diminishing returns, company *A* decided to dismiss one third of its employees, while company *B* drastically raised its prices and installed new fees for previously free services.

We assume that for deciding which company is acting more fairly people use a range of criteria, or *cues*, to evaluate the companies' behavior. For example, a person could ask herself whether a company's decisions harm a large number of people, or whether they damage the environment, or whether they harm a peer or family member. For each company, these cues can take different *cue values*, that is, they can either speak in favor of the company (positive cue value), against the company (negative cue value), or be unknown (neutral cue value). Additionally, cues have different *cue validities*. Cue validities represent the importance a cue has for a person when she is asked to assess a company's decisions. In our study, cue validities were subjective ratings of the cues. From these validities, cues can also be brought into a *cue order*, starting with the cue that is considered most important for judging a company's behavior and ending with the least important one. In our case, this cue order was obtained by asking all participants to rank all cues in the order of their importance for assessing a company's fairness.

We predicted the judgments that participants made using three decision models. These include two *compensatory* models that take all cues into consideration and allow for making trade-offs between different cues by adding all of them up. This means that these models in principle allow for a low value on one cue to be compensated by a higher value on another cue. The two compensatory models are the weighted additive model, here referred to as WADD (sometimes also called *Franklin's Rule*, see Gigerenzer et al., 1999), and a unit weight linear heuristic labeled UWL (for related models see e.g., Dawes & Corrigan, 1974; Gigerenzer & Goldstein, 1996). The third model does not allow for making trade-offs between cues; it is a *noncompensatory* lexicographic decision rule that considers one cue at a time and bases its decision on only one cue (LEX; see e.g., Payne et al., 1993). This model closely resembles the well-known *take-the-best heuristic* (Gigerenzer & Goldstein, 1996) that has been proposed in the fast and frugal heuristics framework as a strategy for making inferences about objective criteria.

Note that we chose these models because they are important representatives of the class of strategies that is studied in the fast and frugal heuristics research program and beyond. However, we do not mean for these strategies to represent the exhaustive set of possible strategies one could envision for moral decision-making. Rather, we see them as a good starting point for investigating the performance of formal models in predicting moral judgments.

Weighted Additive Model: WADD

The WADD model integrates all cues and their validities to determine which of two *alternatives* (here: companies) should be chosen. Specifically, it calculates a weighted *cue sum* for every alternative, and decides for the alternative with the larger cue sum. To compute the cue sum, positive and negative cue values are multiplied by their respective validities:

$$\sum_{i=1}^{n} a_i v_i, \qquad (1)$$

where n is the number of cues, a is the cue value (-1 for negative, 0 for neutral, and 1 for positive), and v is the cue validity.

When two companies are compared, WADD thus proceeds as follows:

(1) For each company, calculate the weighted cue sum.
(2) Decide for the company with the higher cue sum.

Unit weight Linear Heuristic: UWL

Like WADD, the unit weight linear heuristic integrates all cues into a judgment by adding them. It simplifies the task by weighing each cue equally (hence unit weight) and by ignoring cue validities.

It adds all positive and negative cue values for each alternative and decides for the company with the larger cue sum:

$$\sum_{i=i}^{n} a_i, \qquad (3)$$

where a is either -1 (negative cue value) or 1 (positive cue value). The strategy decides as follows:

(1) For each company, compute the sum of negative and positive cues.
(2) Decide for the company with the higher cue sum.

Lexicographic Decision Rule: LEX

Unlike WADD and UWL, this heuristic does not by itself take all cues into account, but operates sequentially and uses only one cue at a time, considering cues in the order of their validity. It bases its decision on the first cue that *discriminates* between the alternatives (companies), that is, the first cue that favors one of them. A cue discriminates between two companies if one of them has a positive cue value and the other has not (i.e., the other cue either has a negative or a neutral cue value). The strategy then favors the company with the positive cue value. This heuristic thus proceeds as follows:

(1) Look up cues in the order of their validity, starting with the cue with the highest validity.
(2) Stop when the first cue is found that discriminates between the companies.
(3) Choose the company that this cue favors.

Experiment

We pitted these models against each other in an experiment. In our study, participants completed two computer-based two-alternative forced choice tasks (henceforth: *choice tasks*) with two types of companies, airlines and automobile companies. Pairs of airlines and automobile companies were created by exhaustively pairing all airlines, and automobile companies, respectively, with each other. For each pair of companies, participants indicated which one they thought made fairer decisions. Subsequently, they completed a computer-based *cue task* in which they evaluated the companies' cue values on a range of cues. Finally, in a questionnaire-based *cue rating task*, they were asked to rank and rate the cues from the cue task in terms of their importance for assessing the fairness of the companies' decisions. We modeled participants' judgments in the choice task using their answers in the cue task and the cue rating task. For each participant, we tested which of the three decision strategies predicted her judgments best.

Method

Participants 40 participants (24 female) were recruited via the subject pool of the Max-Planck-Institute for Human Development and paid 16 € (approx. $21) for participation. Ages ranged from 20 to 35 (mean age 25 years, *SD* = 3.4).

Procedure In the choice task, participants were presented two collections of short newspaper excerpts, each one involving one company. These excerpts described how the company dealt with the current financial crisis.

Specifically, we collected newspaper articles about six airlines and seven automobile companies from the largest German economic newspaper "Handelsblatt" by searching for all articles from the past eight months (January 2008 until July 2008) containing the tags "Luftfahrtbranche" (airline industry) and "Automobilindustrie" (automobile industry). For each company, we then picked those excerpts which most accurately described the decisions that this company made to deal with the current challenges in that industry. We were mostly interested in excerpts that we assumed to be relevant for the reader's moral intuitions about the companies, such as personnel politics, pricing, safety, and environmental issues. To avoid for name recognition and prior knowledge about a company to influence our participants' perceptions of the companies' behavior, we made companies' names anonymous in the excerpts we used for the choice task.

In the choice task, all airlines and automobile companies were exhaustively paired. This yielded 15 comparisons between airlines, and 21 comparisons between automobile companies. In each comparison, for each company the relevant excerpts were shown. The excerpts appeared in random order on the left and right half of a computer screen. Participants were asked to decide which company they thought behaved more fairly by pressing one of the designated keys on the right or left side of the keyboard ('p' and 'q') with their left and right index finger, respectively.

In the subsequent cue task, all newspaper excerpts for each company from the choice task were presented again (for each company the relevant excerpts at a time), and participants were asked to answer 32 questions about each company (e.g., "Could this harm a family member of mine?"). Each question tapped one cue. Participants could answer with 'yes', 'no', or 'don't know' by pressing the designated keys ('p', 'q', and Space). Excerpts for the companies were presented in a random order.

In a questionnaire-based cue rating task, subjects were then asked to rate (between 0 and 100) and to rank (from 1 to 32) the cues from the cue task according to how important they considered them for deciding which company behaved more fairly.

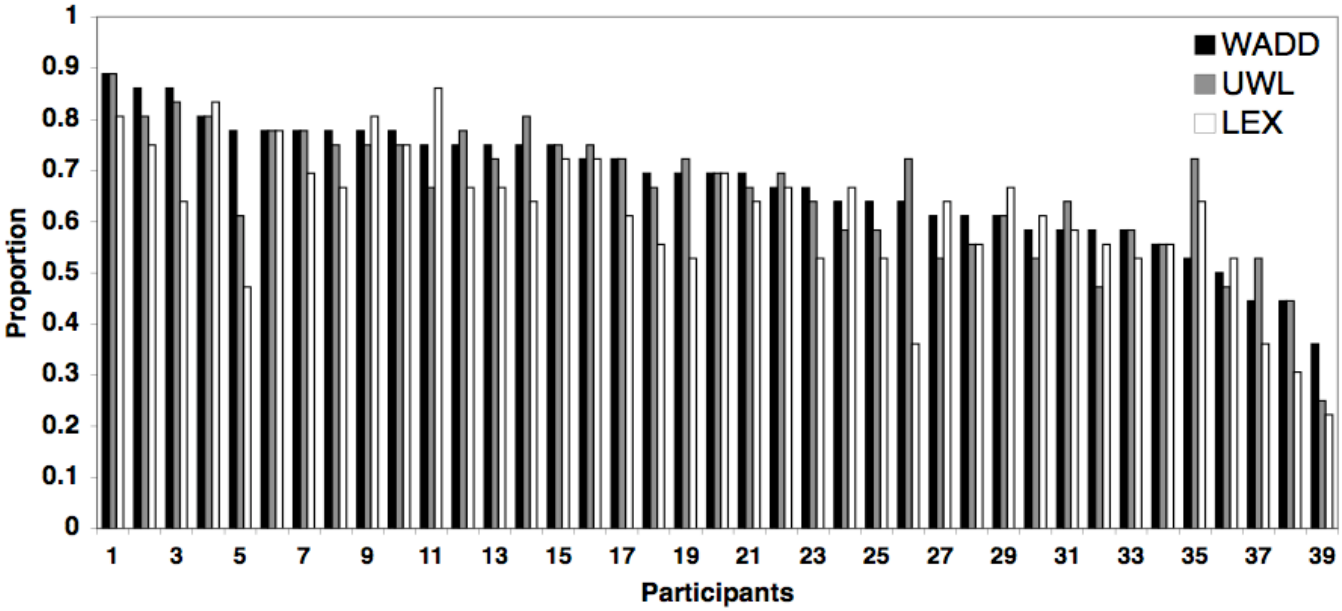

Figure 1: Triplets of bars represent the proportions of correctly predicted decisions, computed for WADD, UWL, and LEX, separately for each participant.

Analyses One subject was excluded from the data set for having submitted a questionnaire with several missing and duplicate rankings and a complete mismatch between cue ratings (validities) and rankings (cue order).

To compare the performance of WADD, UWL, and LEX in predicting participants' decisions, for each model we calculated the proportions of correctly predicted decisions per participant. We will refer to these proportions of decisions made in accordance with a model as a model's *accordance rate.*

To model WADD, we determined the weighted cue sums of each company by multiplying the cue validities from the cue rating task with the cue values obtained from the cue task. To implement UWL, we simply added up all positive and negative cue values for each company as cue sum. LEX was modeled using the cue values from the cue task and the cue order from the cue rating task.

Results

Figure 1 shows the results for all subjects. Triplets of bars represent the proportions of correctly predicted decisions, computed for WADD, LEX, and UWL, separately for each participant. For most subjects, the accordance rates of all three models lies above the chance level of .50; that is, above the proportions of decisions one would expect a model to predict if participants engaged in random guessing. For 13 participants, WADD made the most accurate predictions, 9 participants were best described by UWL, and for 8 participants, LEX proved to be most accurate. Nine participants were equally well described by at least two of the tested strategies.

Discussion

We examined if people's moral decisions can be modeled formally using two compensatory models and one noncompensatory decision heuristic. All tested models could successfully predict a large number of normative judgments made in a two-alternative forced-choice task.

Our results show that normative judgments *can* be predicted with formal models of decision-making. They lend support to the modeling approach to moral decision-making.

Another observation we made is that although WADD made the best predictions for most participants, there were also participants that were best described by UWL or LEX. Such inter-individual differences in strategy use are commonly observed in when studying how people make decisions about objective criteria (see Bröder & Gaissmaier, 2007; Cokely & Kelley, 2009; Mata, Schooler, & Rieskamp, 2007; Pachur, Bröder, & Marewski, 2008). Recently, Feltz and Cokely (2008) could show that individual differences also exist in people's *moral* intuitions when they are confronted with questions about freedom and responsibility. Our findings complement this previous work.

One important caveat of the current modeling is that the number of cues to be rated in terms of their validity was considerably large (32 cues). As a result, participants' cue ratings were most likely prone to some error. One may speculate whether such errors would have hurt the performance of LEX more than that of the other two strategies, which may not depend as much as LEX on that participants' cue ratings mirror exactly their decisions. Moreover, since LEX was the only lexicographic strategy tested in this study, there remains ample room to explore the performance of other noncompensatory strategies for predicting moral judgment.

Conclusion

In this paper, we have taken a step towards formally modeling moral judgments. We showed that one complex decision strategy and two simple decision heuristics could successfully predict the judgments of a majority of

participants. Another finding was that there are inter-individual differences in the strategies that people use to make normative judgments. Further research should focus on testing more decision models for moral judgments and explore how these individual differences emerge.

References

Bergert, F. B., & Nosofsky, R. M. (2007). A response-time approach to comparing generalized rational and take-the-best models of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 107–129.

Bröder, A., & Gaissmaier, W. (2007). Sequential processing of cues in memory-based multi-attribute decisions. *Psychonomic Bulletin & Review*, *14*, 895–900.

Cokely, E.T., & Kelley, C.M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making, 4,* 20-33.

Dawes, R. M. (1979). Robust beauty of improper linear-models in decision-making. *American Psychologist, 34*(7), 571-582.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, *81*, 95–106.

Feltz, A., & Cokely, E. T. (2008). Do judgments about freedom and responsibility depend on who you are? Personality differences in intuitions about compatibilism and incompatibilism. *Consciousness and Cognition, 18*, 342-350.

Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. New York: Viking Press.

Gigerenzer, G. (2008). Moral intuition = fast and frugal heuristics? In W. Sinnott-Armstrong (Ed.), *Moral psychology vol. 2: The cognitive science of morality.* Cambridge: MIT Press.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650–669.

Gigerenzer, G., & Goldstein, D. G. (1999). Fast and frugal heuristics: The adaptive toolbox. In G. Gigerenzer, P. M. Todd, & the ABC Research Group, *Simple heuristics that make us smart.* New York: Oxford University Press.

Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Greene, J. D., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Science, 6,* 517-523.

Greene, J. D., Sommerville, B. R., Nytrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*, 2105-2108.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814-834.

Hauser, M., Cushman, F., Young, L., Jin, R. K., & Mikhail, J. (2007). A dissociation between moral judgments and justifications. *Mind & Language, 22,* 1-21.

Helversen, B. von, & Rieskamp, J. (2008). The mapping model: A cognitive theory of quantitative estimation**.** *Journal of Experimental Psychology: General*, *137*, 73–96.

Marewski, J. N., Galesic, M., & Gigerenzer, G. (2009). Fast and frugal media choices. In T. Hartmann (Ed.), *Media choice: A theoretical and empirical overview* (pp. 107-128). New York & London: Routledge.

Marewski, J. N., & Olsson, H. (2009). Beyond the null ritual: Formal modeling of psychological processes. *Journal of Psychology*, *217*, 49–60.

Mata, R., Schooler, L. J., & Rieskamp, J. (2007). The aging decision maker: Cognitive aging and the adaptive selection of decision strategies. *Psychology and Aging*, *22*, 796–810.

Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Science, 11,* 143-152.

Pachur, T., Bröder, A., & Marewski, J. N. (2008). The recognition heuristic in memory-based inference: Is recognition a non-compensatory cue? *Journal of Behavioral Decision Making*, *21*, 183–210.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York: Cambridge University Press.

Rieskamp, J., & Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica*, *127*, 258–276.

Rieskamp, J., & Otto, P. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*, 207–236.

Sunstein, C. R. (2005). Moral heuristics. *Behavioral and Brain Sciences, 28,* 531-573.