2008

The Moral High Ground: An Experimental Study of Spectator Impartiality

James Konow
Loyola Marymount University, jkonow@lmu.edu

The Moral High Ground: An Experimental Study of Spectator Impartiality

Konow, James

Loyola Marymount University

November 2008

COMMENTS WELCOME November 2008

The Moral High Ground: An Experimental Study of Spectator Impartiality

James Konow
Department of Economics
Loyola Marymount University
One LMU Drive, Suite 4200
Los Angeles, CA 90045-2659
Telephone: (310) 338-2956
Fax: (310) 338-1950
Email: jkonow@lmu.edu

Abstract

Recent developments in various branches of political science underscore the importance of empirical investigation of impartial views. This paper proposes and tests an empirical model of impartiality, inspired by Adam Smith (1759), that is based on the moral judgments of informed third parties (or spectators). The model predicts that spectatorship produces properties widely considered desirable in both normative and descriptive political science research, namely, unbiasedness and consensus. This informs a vignette experiment that elicits moral judgments about real world political economy issues while varying the information conditions and roles of respondents (spectator and stakeholder) across treatments. The results indicate that spectator views are unbiased, and that relevant information reduces stakeholder bias to insignificance, whereas irrelevant information reduces bias but does not eliminate it. Relevant information also promotes a kind of consensus. I argue that this model can complement empirical and theoretical work on deliberation, public opinion and political behavior.

Keywords: impartiality, spectator, consensus, deliberation, public opinion, bias

Ultimately, individuals' beliefs about what is the right or fair economic policy for the nation are difficult to explain. … Since these policy preferences appear to be one of the main forces driving voting behavior, however, explaining them is clearly a key question in American political economy.
– Stephen Ansolabehere, Jonathan Rodden and James M. Snyder, Jr. (2006).

Many questions (about empirical deliberation research) remain, above all concerning the contributions of the individual elements of the deliberative treatment. … How much difference would it make to keep the treatment the same but settle for real-world levels of balance, say?
– James S. Fishkin and Robert C. Luskin (2005).

Can we know, to some degree of certainty, whether moral judgments are genuinely impartial, or, if not, to what degree they are biased by interests or unbalanced information? The first question is critical to the major theoretical traditions in ethics that provide the normative foundation for most political philosophy and public policy research, since those traditions rest on some claim about how actual judgments are, or would be, made under conditions of impartiality. Impartiality and bias also have relevance for political behavior in light of certain recent developments. Many observers of American politics allege that it has become polarized and that political reporting and discourse are characterized by increasing bias, a seemingly inauspicious trend for advocates of deliberative democracy. In addition, recent findings in political science bolster the view that both morals and interests affect issue preferences and that these preferences, in turn, explain vote choice (Ansolabehere, Rodden and Snyder, 2006, 2008). Studying impartiality and bias promises to help us better understand and assess these claims.

This paper proposes and tests an empirical model of impartiality inspired by the spectator theory of Adam Smith (1759). Specifically (and in response to questions raised in the quotes above), this model informs a questionnaire experiment that elicits moral judgments about political economy issues under conditions aimed at reflecting real-world differences in roles and information. Despite a recent explosion of interest in Smith, this is the first detailed model and only thorough empirical analysis based on his impartiality theory. Turning Smith's thought experiment into a laboratory experiment offers numerous advantages, I believe, for addressing a number of interesting questions across broad set of important research agenda in political

science. First, it offers a means to study impartial morals and the biasing effects of interests that figure prominently in public discourse, policy deliberation, political reporting, issue preferences and voting. Second, politics is not solely about "is" but also "ought," and this study belongs to a growing research agenda in political science that explores connections between empirical and normative analysis. At the most general level, the proposed model of impartiality responds to calls for political theory to engage empirical research (e.g., Fung, 2007, and Shapiro, 2002) and to contribute to the empirical investigation of ethics, with its meta-ethical implications (e.g., Mitchell and Tetlock, 2006). More specifically, though, I apply the proposed method to generate evidence on policy preferences in an important and concrete set of policy areas, including the environment, health care and welfare. In this analysis, the focus is on two properties that are often considered key to normative ethics as well as to the healthy functioning of democratic institutions, namely, unbiasedness and consensus.

As I argue below, spectator theory has important advantages over alternate theories as a framework and inspiration for the empirical analysis of impartiality. Nevertheless, I believe this method also complements research on deliberation and public opinion. Thus, although most of this paper treats the spectator approach as it pertains to the relationship between empirical method and normative theory, this approach also relates and potentially contributes to the study of political behavior, public opinion and deliberative democracy, and I will return to these connections in the final section of the paper.

Impartiality can be construed in different ways, so I begin with a review of the main traditions and of the reasons for this paper's focus on the spectator version. Impartiality is central to much of political and moral philosophy, including in some fashion in philosophical systems as diverse as those of Kant, Mill and Nozick, but the three dominant traditions are social contract theory, discourse theory and impartial spectator theory. Social contractarians include Hobbes, Locke and Rousseau, but the best known modern conception of impartiality in this tradition is John Rawls's (1971) original position. The original position is a thought experiment, a hypothetical state in which self-interested individuals initially choose the principles that guide

the basic structure of society behind a "veil of ignorance" of any particulars about themselves, including information about their future position in that society. Rawls held that such a thought experiment leads to unanimous agreement on a rule for organizing the basic structure of social institutions that he called the "difference principle." In a series of novel and informative studies, Frohlich and Oppenheimer (1992, 1987 with Eavy) conducted experiments designed to approximate this kind of impartiality with respect to issues of distributive justice. They found a high level of agreement on trade-offs between multiple distribution rules that spans various countries and cultures. Whereas the studies of Frohlich and Oppenheimer produced little evidence that Rawlsian impartiality leads to Rawls's distributive principle, Michelbach, Scott, Matland and Bornstein (2003) found stronger support for the difference principle but also for the importance of trade-offs between multiple principles.

The link between moral philosophy and empirical social science is inherent to discourse theory (or discourse ethics) as laid out by Jürgen Habermas (1990) [1983]. In this approach, impartiality involves establishing certain presuppositions for argumentation among affected parties. Thus, Habermas's principle of impartiality is, at the same time, a principle for deliberation among stakeholders, viz., one that specifies the conditions for discourse, including equal participation, freedom from constraint, and sincerity. This is summarized in the principle of discourse ethics, which states that "only those norms can claim to be valid that meet (or could meet) with the approval of all affected in their capacity as participants in a practical discourse" (1990, pg. 66). Discourse theory has sparked a sizable volume of political science research, including on "deliberation theory" and on empirical analysis that seeks to relate deliberation theory to practice. This research has been quite varied methodologically. Mansbridge's seminal 1980 book, *Beyond Adversary Democracy*, highlights the challenges to democratic deliberation in the case of two small New England communities. The laboratory experiment of Sulkin and Simon (2001) indicates that deliberation can create more equitable outcomes and enhance fairness perceptions, but these effects depend on the precise implementation. Deliberative polling (e.g., Fishkin, 1991) elicits the views of a random sample of the public after they have been

exposed to balanced information and an opportunity to participate in discussions guided by heterogeneous moderators. List et al. (2007) find that deliberative polls favor a kind of consensus (single-peakedness) and that this effect is related to information gains. Neblo (2007) experimentally tests the links between deliberation and its effects and finds the results mostly consistent with the causality claimed by normative deliberation theory.

Impartial spectator (or ideal observer) theory is mostly closely associated with Adam Smith, although elements of his approach can be traced to Hutcheson and Hume. In his first major work, *The Theory of Moral Sentiments* (or *TMS*, 1759), Smith characterized principles or conditions for judging right conduct, specifically, these conditions are embodied in an informed, sympathetic observer who exists contemporaneously and resides in real people. This impartial spectator utilizes his own experiences as well as his faculties of sympathy as a moral arbiter of his own conduct and of that of others. The interpretation of Smith's spectator implemented in the current study elicits the views of real people, stresses the unbiasedness of third parties relative to stakeholders, and predicts that relevant information helps reduce bias and produce consensus.

The recent, resurgent interest in Smith, particularly in *TMS*, spans philosophy and the social sciences. For example, in political science, Griswold (1999) helps initiate more subtle interpretations of Smith, Hanley (2008) analyzes Smith on building the political and economic institutions of nations, Parrish (2007) sees him as revising moral psychology to make ethical problems of markets seem less problematic, and Rasmussen (2006) offers a solution to the apparent paradox of Smith's support for commerce but admission that money does not buy happiness. In philosophy, Fleischacker (1999) draws on Smith's position on liberty, Fricke and Schütt (2005) provide an extensive set of contributions based on Smith's ethics, and Raphael (2007) gives a careful interpretation of *TMS* close to the one proposed here. In economics, Ashraf, Camerer and Loewenstein (2005) argue that Smith's impartial spectator has broad implications for modern social science research, Konow (2008) examines claims of consensus inspired by Smith's spectator, and Brown (forthcoming) and Nyborg and Brekke (2007) propose a reconciliation of self-interest and moral preferences in a response to the so-called "Adam Smith

problem," which refers to the apparent contradiction between the two. The Haakonssen (2006) volume provides an excellent representation of the breadth and depth of interest in Smith across political theory, philosophy, economics, history and law. Despite the dramatic multi-disciplinary upsurge of research on Smith and *TMS*, the current study represents, to my knowledge, the first thoroughgoing empirical analysis of his impartiality theory, including its examination of the incremental effects of relevant and irrelevant information on moral judgment.

To clarify, the main purpose of this paper is to specify and test an empirical model of spectator impartiality, not to conduct a comparative analysis of possible impartiality theories. Nevertheless, it helps clarify and motivate this study to consider differences and similarities among these three schools of thought. In theoretical terms, different impartiality ideals can easily lead to different optimal rules: for example, a risk loving egalitarian would likely prefer less redistribution as a stakeholder behind a veil of ignorance than as a contemporaneous spectator. That said, the choice of impartiality concept might still be uninteresting, if the theoretical difference is without an empirical distinction, that is, if actual moral judgments do not depend on the version of impartiality implemented. The empirical studies, however, of Amiel, Cowell and Gaertner (2008), and Traub, Seidl, Schmidt and Levati (2005) establish that different impartiality models can produce significantly different moral judgments. I turn, therefore, to a comparison of the theories and the contribution of an empirical spectator analysis.

First, let us note common ground among the three theories. Habermas, Rawls and Smith all associate 1) impartiality with the absence of bias and 2) the satisfaction of the conditions of impartiality with consensus. These facts bolster the use on normative grounds of unbiasedness and consensus as the two principal criteria in our analysis.[1] Discourse theory and spectator theory share a number of common features that simultaneously set them apart from Rawls. For

[1] Habermas, Rawls and Smith are also universalists, who expect their exercises in impartiality to generate general moral rules. This is arguably not a necessary implication of their impartiality theories, since the experiments implied by their approaches to impartiality could conceivably produce weak, multiple or context dependent rules, or, indeed, no rules at all, and some who follow in their footsteps have moved in one or more of these directions.

the former, the relevant moral agent exists contemporaneously rather than in a hypothetical prior state, he is not *solely* self-interested but rather can also access a moral sense, he exists not *only* in idealized form but also in real people who seek to engage the positions and perspectives of other real people, and no information is denied, even about one's own station. There are certain differences, however, in how Habermas and Smith implement this agent. With Habermas, the agent is an implicated party who participates in deliberation with others in order to understand their interests and to moderate his own. For Smith, the ideal agent is a third party, who employs personal experience and imagination to the same ends. In my view, these differences reflect the ethical domains that are foremost in the minds of their architects: Smith's applications are mostly to social interactions and institutions that are on a smaller scale, whereas Habermas focuses on a larger scale effects. These differences fit the contexts they originally envisioned.[2]

The shared features of spectator and discourse theory give those approaches significant practical advantages over contractarian theories for the purposes of empirical analysis: their moral arbiters are real, contemporaneous, endowed with a moral sense and not denied information about themselves and their experiences. Among theories of impartiality, however, spectator theory is unique in adopting a third party perspective, as opposed to that of actual or hypothetical stakeholders. Moreover, spectator theory has not previously undergone a detailed empirical examination, in contrast to theories in the Rawlsian and Habermasian traditions. The spectator approach is, I believe, not only intuitively appealing but also provides a compelling method for shedding light on important open questions in political science, including on the role of moral views and interests in issue preferences and how to identify balanced information in deliberative processes. This is not to suggest, however, that the spectator approach replace other methods. Indeed, I believe spectator, deliberation and public opinion research are potentially

[2] For example, in deciding how much time I should volunteer for my child's school or a political campaign, objective reflection about uses of my time and the impact of those uses on others has various advantages over convening a committee to deliberate the issue. On the other hand, in selecting a tax structure, the participation and buy-in of affected parties is potentially important both to constructing and to enforcing such a system.

complementary and can enhance and strengthen one another in practice. The concluding section contains thoughts about this and about other potential contributions of spectator theory.

The paper is organized as follows. The following section describes the empirical spectator model in general form. The section Design and Procedures presents the specific application of the spectator model in this study, which involves a questionnaire-experiment. The Results and Analysis section summarizes and analyzes the findings with respect to bias and consensus. The final section discusses the results, particularly as they relate to other theoretical and empirical work in political science.

THE "QUASI-SPECTATOR" MODEL OF IMPARTIALITY

This section presents an empirical model of impartiality, including a review of how it takes its inspiration from Smith's *TMS* and of empirical tests and hypotheses that extend that work. The chief contribution of *TMS*, in my view, is not to ethics (understood broadly as the study of moral conduct) but rather more specifically to moral epistemology; that is, this work proposes conditions for acquiring moral knowledge. That Smith's aim was epistemological is suggested by the sub-title of *TMS*: "An essay towards an analysis of the principles by which men naturally judge concerning the conduct and character, first of their neighbors, and afterwards of themselves." This sub-title also portends the dual nature of the agent Smith proceeds to develop as both observer and participant. The "supposed" or "ideal" spectator is a disinterested (or rather, *detached*) third party, whereas the "real" spectator attempts to apply reflexively the lessons of the former: "We endeavour to examine our own conduct as we imagine any other fair and impartial spectator would examine it." Both agents figure prominently in Smith's writing, and he distinguishes degrees of spectatorship from the imperfections of a child or person of weak character to the self-command of the ideal spectator.

This take on spectatorship also implies a more nuanced view of sympathy in *TMS*. Sympathy involves adopting the positions of others, both cognitively and affectively. I see this as serving two functions in Smith, similar to the distinction Rawls (2000) makes in reference to

Hume's spectator. On the one hand, sympathy has an *epistemic* role that is chiefly of relevance to the ideal spectator: it enlarges his awareness of relevant facts by enabling him to factor in the experiences and feelings of others in coming to moral judgments. On the other hand, sympathy has a *motivational* function that pertains to the real spectator and helps him to put aside, or at least to moderate, his own interests relative to those of others.

Smith states that the relevant judgments are those of the "impartial and well-informed spectator." This third party is not an unattainable abstraction but rather a state that real people can achieve, at least to an approximation. Importantly, this spectator accesses his own life experiences: "The man who is conscious to himself that he has exactly observed those measures of conduct which experience informs him are generally agreeable, reflects with satisfaction on the propriety of his own behaviour." Nevertheless, his conduct is not motivated by a desire for social approbation or by false consciousness, since "he views it in the light in which the impartial spectator would view it, … and though mankind should never be acquainted with what he has done, he regards himself, not so much according to the light in which they actually regard him, as according to that in which they would regard him if they were better informed."

These passages help underscore the emphasis of this study on two qualities of the ideal (but not real) spectator. First, the spectator is not and never will be implicated in a situation being evaluated, that is, he has no stake, real or perceived, that might bias his judgments of right and wrong. Second, the spectator seeks to be fully informed of the relevant particulars and processes this information rationally with respect to internalized values. Since the focus here is on the ideal spectator, sympathy can be subsumed under this second point in which it serves an epistemic, or informational, function.

Spectatorship is not merely a philosophical ideal but also a standard we can recognize in a variety of social institutions and practices. In the justice system, the selection of juries, the assignment of judges and the rules of evidence are designed to approach third party impartiality while liberally providing relevant information. Politicians are subject to conflict of interest provisions, and an important impetus to campaign reform is to sever legislators from stakes that

might bias how they represent the interests of their constituencies. Although its efficacy is debatable, government regulation, at least as an ideal, provides informed, third party oversight of industries. Sometimes even forces in the private sector attempt to bolster their impartiality credentials by supporting informed (ostensibly independent) verification of their practices, such as the Better Business Bureau or ClimateCounts.org, a privately supported organization that reports the performance of various businesses to limit their greenhouse gas emissions.

Of course, the ideal conditions of Smith's impartial spectator probably never obtain in the real world. Information is incomplete, people interject their own interests, and even judgments of otherwise objective persons can be tainted by unrepresentative experiences and beliefs. On the other hand, the phenomena described above, which reflect the ideal, also demonstrate its normative value. Moreover, the approach adopted in the current study explores spectatorship under considerably more controlled conditions and with the more focused goals of informing and of assisting reflection on theory and policy. Nevertheless, complete information and the utter absence of stakes are presumably never both achieved in the real world. Given this, can spectatorship be implemented empirically, and, if so, how?

The method proposed here seeks to answer the above question by examining the judgments of a "quasi-spectator," whose views are taken as an approximation of those of the ideal spectator. The quasi-spectator is a real world observer who evaluates the moral rightness of a situation in which he has no salient stakes but about which he is well, albeit probably never completely, informed. The proposed method uses conditions designed to elicit and analyze the impartial moral preferences of real individuals. The general conditions are quite simple and relate to *role* and *information*. First, subjects should not be salient stakeholders in the matter they are judging, and the elicitation of their preferences should distance them as much as possible from their own interests. Thus, the expression of preferences should be separated, to the extent possible, from any rewards or sanctions. Ideally, spectator compensation, if utilized, is unrelated to decisions they make about others, and spectators are not known personally to or by other subjects in order to avoid activating non-ethical (e.g., social approval) motives. Second,

information should be carefully chosen so as to provide vital context that allows subjects to respond with respect to their moral preferences but with an awareness of the danger of priming personal bias. Contextual details can facilitate moral reasoning, as they have been shown to do in vignette studies (see, for example, Goldstein and Weber, 1995), but the researcher should avoid, or at least control for, specifics that could result in a self-interest bias, e.g., through material stakes or vicarious identification with one party over another.

These conditions can be implemented in a variety of ways using different empirical methods, including with vignettes, attitude surveys and paid laboratory and field experiments.[3] Indeed, many studies across the social sciences (including public opinion surveys) and in experimental philosophy have employed quasi-spectators in one form or another. But what evidence is there that quasi-spectator views even approximate those of the ideal spectator? What distinguishes the current study is its proposal to answer this question by stating and testing certain empirical conditions. The *quasi-spectator model* refers to variation in role and information to examine the effects on two properties that have wide appeal on normative and empirical grounds: *unbiasedness* and *consensus*. This is elaborated below.

Absence of bias is universally associated with impartiality. The version of unbiasedness adopted here is in the long Aristotelian tradition of virtue as a mean: evidence of quasi-spectator unbiasedness is present if their views are a (not necessarily arithmetic) mean of the views of stakeholders whose interests are in opposition. On the second property, this study takes seriously the claim of most prescriptive theories that ideal impartiality results in unanimity and extends this reasoning by arguing that real world impartiality might not produce unanimity, but it should result in *consensus*. Here this is defined as a convergent trend of opinion by quasi-spectators as information relevant to their values is added, and, hence, will also be referred to as *convergence*. Thus, in this paper, consensus does not necessarily involve unanimity but rather a higher level of

[3] I should note some differences on a matter of terminology as they pertain to the project reported here. Since this study involves manipulation of independent variables and random assignment of subjects to treatments, most psychologists, sociologists, philosophers and political scientists would characterize it as an experiment, but most economists would not, since subjects are not compensated in monetary terms.

agreement. The rationale for consensus is that relevant information provides the means for quasi-spectators to reason from a common set of values and reduces extraneous influences. If consensus did not emerge, that would be consistent with some failure of the quasi-spectator method, e.g., it could indicate that individuals do not share the same underlying moral views and information enables them to judge from these heterogeneous values, or it could indicate insurmountable obstacles to practical objectivity, e.g., if information merely feeds self-interested biases. Thus, even if one accepts the impartial spectator in principle, there remains this empirical claim that any implementation based on it must establish, which the current study addresses.

Unbiasedness and consensus enjoy widespread acceptance in the three dominant schools of thought about impartiality, but one should acknowledge that there are dissenting views about the normative appeal of consensus. Radical pluralists (e.g., Mouffe, 1996) and difference democrats (e.g., Young, 1996) oppose consensus, highlighting practical difficulties with procedures to achieve consensus and arguing that the good is plural. The finer points of this thinking go beyond the scope of this paper, but these critiques have contributed to a shift away from unanimity and toward broader concepts, including *meta-consensus* (Dryzek and Niemeyer, 2006) and *single-peakedness* (List et al., 2007). The convergence proposed in this paper also represents a broadened concept of consensus, indeed, one that does not require that one abandon (or, for that matter, embrace) unanimity as a theoretical ideal.

Spectator theory, as formulated in *TMS*, has certain implications about how role and relevant information affect bias and consensus, but it does not address all role/information possibilities, e.g., the effects of irrelevant information, so I distinguish below *core hypotheses* from supplemental *conjectures*. Consider first the four core hypotheses of the quasi-spectator model. Note that the term "spectator" will henceforth be used as shorthand for "quasi-spectator," whose views approximate those of ideal spectators, and the term "stakeholder" will refer to an individual whose moral views are potentially influenced both by his own stakes in the matter at hand as well as by the judgment of the impartial spectator, as with Smith's "real" spectator.

H1. *Stakeholders are biased, at least in the absence of sufficient relevant information, and*

spectator judgments are intermediate to opposing stakeholder views.

At least under low information conditions, stakeholders are biased in the direction of their opposing interests and away from impartial views, so spectator judgments are intermediate.

H2. *Relevant information creates consensus among spectators.*

Relevant information permits spectators to judge more accurately from common moral views.

H3. *Relevant information reduces bias among stakeholders.*

Stakeholder views partially reflect the sentiments of spectators. Thus, they should respond to relevant information by attraction to the shared, and now more identifiable, moral position.

H4. *Relevant information creates consensus among stakeholders.*

Again, relevant information should have similar effects on stakeholders as spectators, since spectators are embedded in stakeholders.

Spectator theory does not speak to the potential effects of morally *irrelevant* information. Nevertheless, in the less than ideal circumstances of the real world, both relevant and irrelevant information are often present and difficult to disentangle.[4] These data, however, allow one to explore whether and how different combinations of relevant and irrelevant information affect average views and consensus. If relevant and irrelevant information have different effects on bias and consensus (e.g., relevant information has desirable effects whereas irrelevant information has undesirable effects), that would recommend distinguishing the two types of information, as a practical matter. Moreover, these information effects could differ for spectators and stakeholders. Theoretical work in philosophy (e.g., Rawls, 1971) as well as empirical social science research (e.g., Babcock and Loewenstein, 1997, Dunning, Meyerowitz and Holzberg, 1989) suggest that information provides the raw material for stakeholder bias. On the other hand, information that increases spectator consensus might do so for stakeholders, as well. Spectator theory does not address these important questions, but I propose two conjectures, **H5** and **H6**, about them. Since arguments regarding the effects of irrelevant information go both ways, the null hypothesis is

[4] For now, the distinction between relevant and irrelevant information is not explicitly defined. In the following section, however, more precise definitions will be stated in the context of this study.

that irrelevant information has no effect, regardless of presence of relevant information.

H5. *Irrelevant information does not significantly affect the moral views of stakeholders, whether or not it appears in combination with relevant information.*

This conjecture is silent about spectators for reasons explained in the following section.

H6. *Irrelevant information does not significantly affect spectator or stakeholder consensus, whether or not it appears in combination with relevant information.*

Having laid out the quasi-spectator model of impartiality in general terms, we now turn in the following section to the specifics of how it is implemented in the current study.

DESIGN AND PROCEDURES

This study employs questionnaires that ask respondents to allocate fixed resources in six hypothetical scenarios (or vignettes) under different conditions with respect to their role (stakeholder, spectator) and available information (base, relevant, irrelevant). Many previous studies of moral intuition have employed *contrasting* versions of scenarios (e.g., Kahneman, Knetsch and Thaler, 1986), but the quasi-spectator model requires examining the *incremental* effects of information on moral judgment, which only a few studies, such as Faravelli (2007) and Konow (2008), have utilized. Moreover, the latter two studies involved only two conditions per scenario, as opposed to the current study, which employs twelve conditions per scenario and is unique, to my knowledge, in considering the effects of role and the quantity of relevant and irrelevant information on bias and consensus.

The goals in designing the study were to elicit thoughtful and explicitly moral judgments to situations in which information is carefully controlled. Moreover, the design is unusually ambitious for studies of this kind with 72 total treatments (six scenarios and twelve role/information conditions per scenario). These facts informed the choices regarding method and subject pool.[5] A comparatively large number of observations was needed given various

[5] The basic empirical method and subject pool for this study correspond to the practice in many political science experiments and in almost all studies in the field of experimental philosophy (see, for example, Knobe 2007, Nadelhoffer and Nahmias 2007, and Nichols 2004), including what Mitchell and Tetlock (2006) call experimental political philosophy.

considerations, including the numerous versions of each question, the fact that individual subject questionnaires did not usually contain the full complement of scenarios, the collection of a minimum of 50 observations for each version of every question (and frequently more than 100), and the use of a between subjects design for role and information. These facts alone necessitated more than 1000 subjects, but, in addition, revisions to vignettes involved multiple waves that brought the actual total number of subjects to 1383. Participants, therefore, were drawn from undergraduate students enrolled in social science classes at a comprehensive US university who participated in order to complete a course requirement. Although the subjects were disproportionately lower classmen, they were otherwise quite representative of students at this demographically diverse institution, as illustrated in Table A1 of the Appendix.[6]

TABLE 1. Summary of Contexts and Vignettes

Question	Ethics Field/Subject	Policy Area
1	Environmental ethics	Environmental protection
2	Law and ethics	Tort law
3	Bioethics	Health care
4	Accountability principle	Labor compensation
5	Need principle	Welfare
6	Efficiency principle	Industrial restructuring

The six scenarios in the questionnaire are classified in Table 1 according to the ethics field or subject they reference and the particular policy area they address. All scenarios involve economic policies or issues with significant economic content for two reasons: on the one hand, the design requires quantitative dependent variables, and, on the other hand, recent research underscores the importance of such issues for public opinion and political behavior (Ansolabehere et al., 2006, 2008). Questions 1 through 3 deal with fields in applied ethics,

[6] To the extent this sample differs from the more general population, that is probably not critical to this study, since its focus is not on subject views per se but rather on the relative effects of allocator role and information on those views, and there is no obvious reason why or evidence that these effects would be specific to any group (indeed, see Tables A4 to A6 of the Appendix demonstrate the almost universal insignificance of demographic variables for subject responses). The data were collected from November 2003 to April 2006.

whereas questions 4 through 6 are designed to represent contexts in which a single distributive principle is salient using the justice framework proposed in Konow (2003). Specifically, question 1 concerns the amount by which the EPA should require a pulp mill to reduce the waste it discharges into a river in light of the consequences for workers at the mill as well as those who live near the river. Question 2 asks how much the court should award the victim of a vehicular accident considering the damages and responsibilities of the two parties.[7] Question 3 asks how the limited budget of a hospital should be allocated given the need to provide both emergency services and preventative services to the local community. Question 4 addresses worker compensation when the salient fairness concern is accountability (Konow, 2000), i.e., rewarding agents in proportion to the contributions they control (here work hours). Question 5 focuses on need-based government support for a student trying to complete a high school diploma. Question 6 concerns the efficient use of limited investment funds to accommodate industrial restructuring given changes in technology and industry demand.

TABLE 2. Experimental Design and Treatment Acronyms

	Information condition			
Allocator role	Base	Base+Relevant	Base+ Irrelevant	Base+Relevant +Irrelevant
Spectator	S	SR	SI	SRI
High stakeholder	H	HR	HI	HRI
Low stakeholder	L	LR	LI	LRI

There are twelve versions of each of the six questions, which vary according to the information provided and the role into which the respondent is cast. This is illustrated in Table 2. Respondents are instructed to select the amount they think "should" be allocated to some party or purpose, where permissible choices are on a closed interval that, for ease of comparison across questions, runs from zero to a power of ten (i.e., 10, 100, 1000, etc.). For each question, a respondent allocates from one of three possible hypothetical roles: an uninvolved third party or

[7] This scenario is closely modeled on the case used in a series of studies reported in Babcock and Loewenstein (1997). I wish to thank Linda Babcock for kindly sharing their materials.

Spectator, a person who benefits directly or by association from having an amount that lies on the lower (i.e., zero) end of the allocation interval or *Low Stakeholder*, or a person who prefers an amount that lies on the upper end of the allocation interval or *High Stakeholder*. Note that Low and High do not mean small and large stakes, respectively, but rather indicate the direction of preference along the interval. For example, in question 1, the Spectator is told the pulp mill is in a different region of the country so as to lessen any projected concern for one's own employment or exposure to pollution, the High Stakeholder (who wants a large reduction in emissions) lives near the mill, and the Low Stakeholder (who wants a small reduction in emissions) commutes from elsewhere to work at the mill. These three roles are crossed with four information conditions: the Base information, which is identical in all conditions, the Base plus Relevant information, the Base plus Irrelevant information, and the Base plus both Relevant and Irrelevant information. Table 2 also lists the acronyms that are used for these treatments. Thus, there are 72 conditions in this six Scenario (questions 1–6) × three Allocator role (S, H, L) × four Information condition (Base, R, I, R+I) factorial design. All subjects responded to more than one scenario, but the role and information conditions were between subjects: no respondent read more than one of the twelve versions of each question. The text of the twelve versions of the six questions can be found in composite form in the Appendix.

Although it is usually obvious whether a role in these contexts is that of a Spectator (S), High Stakeholder (H) or Low Stakeholder (L), it is less clear whether information is morally Relevant or Irrelevant or what those terms should mean. I chose an empirical definition for these concepts inspired by spectator theory. Information is considered morally Relevant (R) if it causes a statistically significant shift in the mean judgment of spectators when added to their base information (i.e., the means of S and SR differ). For example, for question 1, information about the consequences of different levels of pollution reduction on nearby residents and on workers at the pulp mill is relevant. Information is deemed morally Irrelevant (I) if, when added to R information, it does not produce a statistically significant shift in mean judgments of spectators (i.e., the means of SR and SRI do not differ). For example, for question 1, information about the

opposing positions of the labor union and local homeowners' association is irrelevant. That is, R information causes impartial third parties to change their moral views, whereas I information does not. To create substantial separation between these two types of information, in this study information must cause a shift in means at the 0.1% level of significance to be considered R and must produce no difference in means even at the 10% level of significance to be considered I. Although the content of most scenarios was chosen ex ante, identifying passages that qualified as R or I according to these criteria sometimes involved revision of content. Since statistical criteria are used to define and identify R and I, statistically testing differences based on such information might seem circular, but that is not so: these difference in means tests of R and I help establish independent variables, but the hypotheses to be tested employ a different set of dependent variables on bias and consensus, to which we now turn.[8]

Bias is measured in this study by comparing mean decisions under different conditions. If the manipulation is effective, under low information conditions the mean H view should exceed the mean L view, whereas unbiased spectator views will lie between these opposing stakeholder views (**H1**). Consensus is identified by comparing variances in views and is associated with a decrease in variance across conditions, e.g., with the addition of R (**H2** and **H4**). It should be noted that, in the vignettes presented in this study, the first two moments of the distribution of views can be used to address bias and consensus. It is not, however, the general claim of the quasi-spectator model that these two metrics are always the correct ones for these purposes.[9]

Numerous measures were undertaken to achieve good survey design. In order to promote respondent attentiveness, subjects never answered more than six questions and often answered fewer, depending in part on the overall length of their questionnaires. For similar reasons, long

[8] This is the reason spectators are not addressed in **H5** in the prior section: their moral views are used for identifying R and I.

[9] For example, suppose all spectators agree completely on a rule that allocates rewards, y, as a function, $f(x)$, of individual trait x, but different spectators observe individuals with different values of trait x. Then, the ideal distribution of judgments is not degenerate. Unbiasedness means here that the expected deviation of y from its ideal (given x) is zero, while consensus means lower variance of judgments controlling for morally relevant parameters such as x.

versions of questions were generally balanced with short versions of other questions. Simple and clear instructions directed subjects to choose the allocation they thought "should" be enacted in each scenario, i.e., even stakeholders were asked to choose the morally right allocation. To address possible order effects, a randomized Latin square design was adopted, i.e., versions of questions were randomly assigned to a variety of different orders. Demographic data were collected from subjects on age, gender, race, and a variety of academic and economic variables.

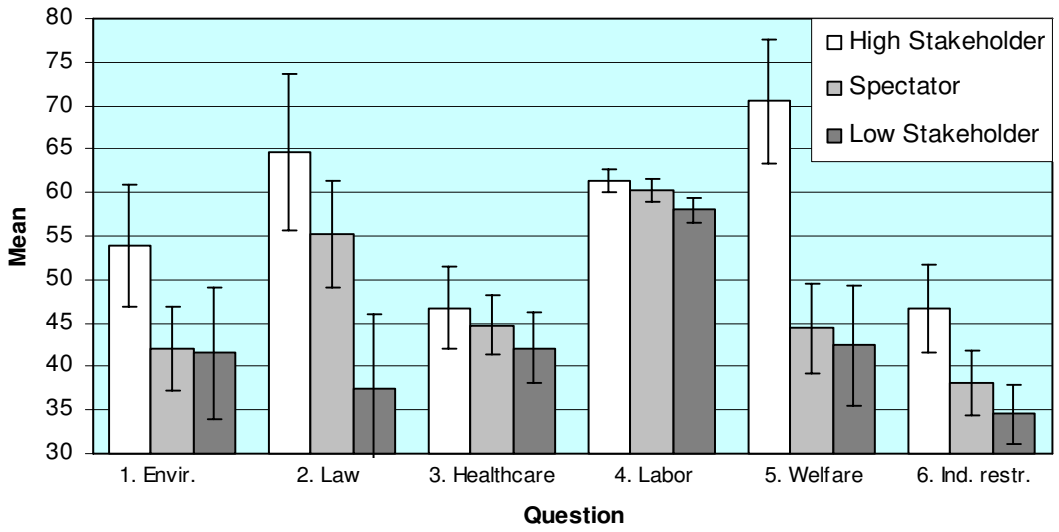
RESULTS AND ANALYSIS

The presentation and analysis of the results in this section focus on comparisons of means and variances between different allocator role and information conditions. In order to facilitate comparisons across different questions, the responses have been converted, where necessary, to a common 0 to 100 scale (dividing by 1000 for question 2, dividing by 10 for question 5 and multiplying by 10 for question 6). Table A2 in the Appendix reports the results of tests that confirm that the information used in the questions comply with the criteria defined for them. Specifically, according to two-tail t-tests of differences in mean judgments, R information, when added to the Base, causes a shift in the mean judgments of S that is significant at the 0.1% level, whereas I information, when added to R, produces no difference in judgments that is significant at the 10% level (indeed, there is only one such shift that is significant even at the 20% level).

Figure 1 summarizes the mean responses by allocator role for each of the six questions under the Base information condition with bars, while the 95% confidence intervals are indicated with lines. The mean judgments of H always exceed those of L, indeed, these differences are significant at the 1% level, except for question 3 where p=.07 (one-tail t-tests). This systematic pattern of bias is particularly striking, in light of the fact that even stakeholders were asked to provide their moral preferences (what they think *should* be implemented), which one would expect to be less biased than their personal preferences (what they prefer to see enacted). In addition, the mean Spectator responses consistently lie between the means of their respective H and L responses for all questions. These results corroborate the effectiveness of the manipulation

of roles and are consistent with the predictions of **H1**.

Figure 1. Mean Responses by Role and Question
(Base Information Condition)

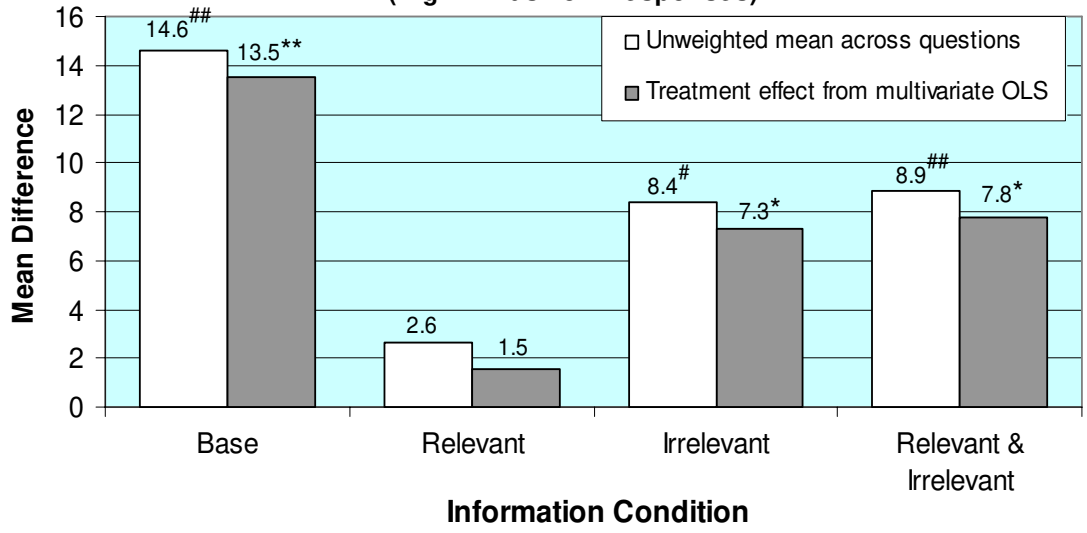


Note: Lines indicate the boundaries of the 95% confidence intervals.

What are the effects of relevant and irrelevant information on stakeholder bias? Figure 2 illustrates average stakeholder bias under the four information conditions, where bias is measured as the difference between responses of high and low stakeholders. The white bars represent the unweighted mean bias across the six questions, where # indicates bias is significant for 3-4 questions and ## for 5-6 questions. The dark bars are the estimated effects from a regression of average bias for each of the six questions and four information conditions on dummy variables for the questions and information conditions, where question 1 and R are the omitted categories (see column (1) of Table A3 in the Appendix). As is apparent, both measures suggest quantitatively similar effects. With only Base information, there is, on average, approximately a 14 point stakeholder bias. When relevant information is added, this bias is reduced to insignificance, strongly supporting **H3**. Interestingly, irrelevant information alone also reduces bias in comparison to the Base condition (F-statistic=6.17, p=.03), but not to insignificance (t-statistic=2.30, p=.04). Even more interesting, perhaps, is the fact that bias

equals roughly the same 8 percentage points with I as with both R and I: in the latter case, bias is also significantly lower than in the Base condition (F-statistic=5.19, p=.04) but significantly greater than zero (t-statistic=2.50, p=.02). Thus, these results refute the conjecture **H5** that irrelevant information has no effect on bias and have more subtle implications: both R and I information reduce bias, but eliminating significant stakeholder bias altogether involves providing R while withholding I.

Figure 2. Stakeholder Bias
(High minus Low Responses)



Notes: #/## indicates that 3-4/5-6 of the six questions differ significantly from zero at the 5% level; */** means the regression coefficient differs from that in the Relevant information condition at the 5%/1% level of significance (two-tail t-tests).

Table 3 summarizes the results of regression analyses that break down the effects of information on moral judgments by allocator role and scenario. These regressions also include controls for the personal characteristics of gender, race, major, class, age, expenditures, parents' income, hours worked and earnings, which are reported in the Appendix. This analysis allows us to examine the information effects by question and allocator role, to address the significance of R and I controlling for personal characteristics, and to explore possible effects of personal characteristics that might reflect personal bias. Relevant information (R and R+I) drives judgments significantly and in the same direction for spectators and stakeholders in 97% of the

cases (all 36 cases except R for L on question 6). Five of the I dummies are now significant, but the effects of I remain smaller in magnitude than those of R (with one exception) and, controlling for personal characteristics, remain insignificant in 72% of the cases. The effects of the eleven personal characteristics are mostly small in magnitude and usually not significant: only 6% (twelve) of these 198 coefficients (eleven variables for three allocator roles and six questions) are significant at conventional levels, roughly the percentage one would expect by chance (see Tables A4 to A6 in the Appendix).

TABLE 3. Regression Analysis: Moral Judgments of Spectators and Stakeholders

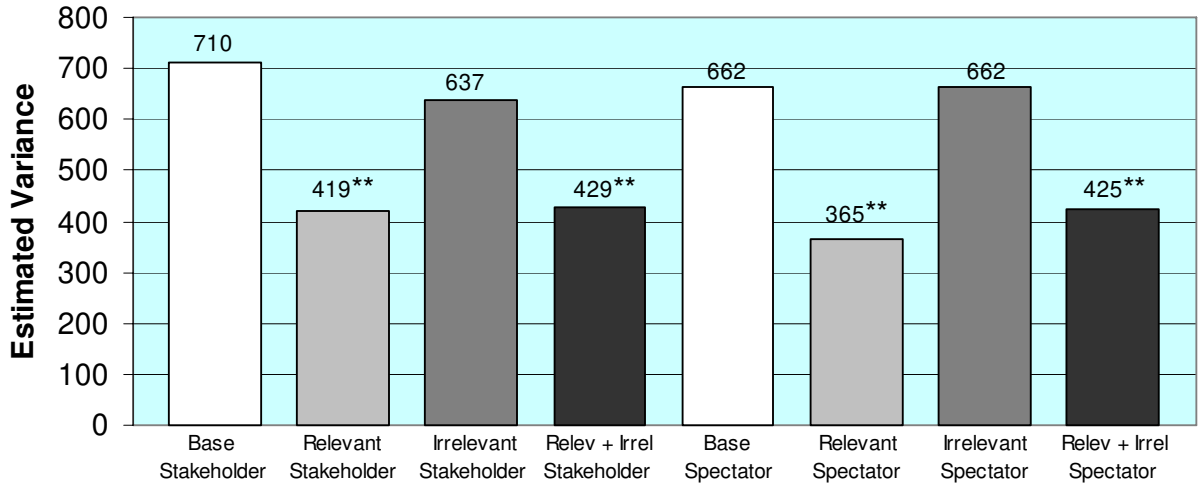
	Question					
Regressors	1. Environ.	2. Law	3. Bioethics	4. Accountab.	5. Need	6. Efficiency
Spectators						
Relevant	18.96**	−21.71**	14.14**	13.37**	32.95**	−12.07**
Relev+Irrelev	16.63**	−17.38**	14.18**	12.23**	27.87**	−10.51**
Irrelevant	6.97^	0.62	−2.21	−1.30	9.31*	−8.13**
Observations	309	332	314	327	310	324
R-squared	0.15	0.13	0.25	0.59	0.31	0.13
High stakeholders						
Relevant	10.37*	−35.59**	9.97**	10.64**	14.14**	−14.91**
Relev+Irrelev	12.10**	−27.00**	15.69**	9.60**	7.82*	−15.33**
Irrelevant	0.17	12.69*	−0.52	0.85	−6.15	−8.00*
Observations	225	219	226	223	221	223
R-squared	0.13	0.12	0.22	0.33	0.16	0.24
Low stakeholders						
Relevant	19.34**	−15.43**	16.71**	10.27**	32.40**	1.06
Relev+Irrelev	16.71**	−15.88**	13.53**	9.50**	23.06**	−11.68**
Irrelevant	5.87	16.08**	−0.01	−0.35	8.73^	−0.02
Observations	224	216	238	220	215	227
R-squared	0.11	0.28	0.25	0.21	0.26	0.14

Note: The entries are regression coefficients. The full results of these regressions are reported in the Appendix and include controls for gender, race, major, class, age, expenditures, parents' income, hours worked and earnings.
^p<.10, *p<.05, **p<.01 (two-tailed)

Turning now to the matter of consensus, Figure 3 presents the results of a regression of variance on allocator roles and information conditions, controlling for questions (see column (2)

of Table A3 in the Appendix). The bars represent the estimated variance under different role and information conditions, where asterisks indicate the significance of differences in the respective regression coefficients from the Base Stakeholder condition (stakeholders include both H and L). What stands out is that estimated variance is consistently lower with R than without R, indeed, by more than 40%, on average. Given the presence (or absence) of R, variance does not differ with allocator role or I information even at the 40% level of significance, but all conditions with R differ from all conditions without R at the 5% level (see Table A7 in the Appendix). Thus, relevant information is decisive: it reduces variance significantly and by a similar magnitude that does not depend on whether or not irrelevant information is present. Irrelevant information does not significantly affect variance. These results corroborate all predictions about consensus: relevant information creates consensus among spectators (**H2**) and stakeholders (**H4**), and irrelevant information does not significantly affect spectator or stakeholder consensus (**H6**).

Figure 3. Effects of Role and Information on Consensus



Notes: */** indicates that the estimated variance differs at the 5%/1% level of significance from that in the Base Stakeholder condition (two-tail t-tests).

DISCUSSION AND CONCLUSIONS

This paper proposes a model of spectator impartiality inspired by Smith, tests it

empirically, and finds significant support for its core hypotheses (**H1** to **H4**) and interesting results on two related conjectures (**H5** and **H6**). In the absence of relevant information, stakeholders are biased (**H1**), but relevant information reduces this bias (**H3**). Irrelevant information also reduces bias (contrary to **H5**), but eliminating bias appears to involve providing relevant information without irrelevant information. Relevant information creates consensus (reduces variance) among spectators (**H2**) and stakeholders (**H4**), whereas irrelevant information does not significantly impact spectator or stakeholder consensus (consistent with **H6**). I conclude with some reflections on spectator theory and how it might complement future work on deliberation, public opinion and political behavior.

The ideal spectator described in Smith's *TMS* need only introspect to know what is right, since such agents have shared values that lead to the same conclusions. A practical weakness of this approach is its reliance on a single agent: even if the presupposition of shared values is correct, one cannot expect the real world to display the conditions necessary for unanimous, unbiased judgments. The quasi-spectator model addresses this by accessing the moral intuitions of multiple individuals and by a broader concept of consensus, similar to developments in deliberation theory. It takes unbiasedness and consensus as necessary and sufficient conditions for impartiality. As critics of consensus point out, consensus is not desirable if there is bias. For instance, consensus is normatively suspect, if it is achieved through coercion by the more powerful, conformity due to differing skills of persuasion, framing effects or false consciousness. Similarly, the value of unbiasedness depends, in this framework, on some level of consensus. In instrumental terms, moral views that are unbiased but more dispersed can contribute to higher rates of dispute that prove costly to all parties, consistent with experimental results on dispersion and bargaining (e.g., Babcock and Loewenstein, 1997). In intrinsic terms, consensus is a logical consequence of unbiasedness itself for universalists like Habermas, Rawls, Smith and their followers.

I believe the empirical methods of deliberation, public opinion and the quasi-spectator provide lessons that can prove mutually beneficial, including for the analysis of political

behavior. Deliberative polling, such as Fishkin and Luskin (2005), treats the benefits of combining the first two methods, and Barabas (2004) proposes a theory of how deliberation increases knowledge and alters opinions, building in large part on public opinion research, especially Zaller and Feldman (1992). I will focus on the other two combinations of methods.

Public opinion research typically involves mass surveys on policy issues whereas the quasi-spectator method specifies general conditions for eliciting moral preferences, which can include surveys or other methods. Their intersection constitutes large sample surveys that explore moral preferences over public sector issues with an awareness of role and/or information conditions. A nice example along these lines is Gibson's study (2008), which embeds an experimental vignette in a representative survey to examine perceptions of judicial impartiality. Within this research area, the quasi-spectator method suggests possible measures to implement impartiality and evaluate bias. For example, respondents might be asked for moral preferences (what policies they think *should* be enacted) as well as personal preferences (what policies they prefer). Bias can further be studied by tracking variables related to participant stakes in the issues they are judging, and perhaps by consciously selecting respondents for their second party or third party status. Public opinion research, in turn, provides a wealth of lessons for quasi-spectator studies that use surveys, including about subject pool selection, salience, order effects, wording effects, framing effects, interviewer effects, reference group effects, consistency, and reliability.

This study has focused chiefly on the normative value of identifying impartial moral judgments and stakeholder biases, but recent findings also suggest possible benefits of the spectator method for descriptive research on political behavior. Ansolabehere, Rodden and Snyder (2006, 2008) identify well behaved issue preferences that have significant explanatory power for voting behavior and find that, even on economic issues, a large part of the variance in these preferences cannot be traced to self-interest.[10] They conclude that understanding beliefs

[10] I should note an important difference in the use of the term "moral." In the Ansolabehere et al. articles, "moral" issues include views of abortion, gay rights, religion and school prayer, as opposed to economic issues, which include views on the desirable scale of government and its involvement in the economy. In

about what is right or fair is a difficult but also key question, since these are an important determinant of voting behavior (2006). The quasi-spectator method offers a means for eliciting such beliefs; this, in turn, can enable efforts to construct theories of beliefs about what policies are right or fair (e.g., Konow, 2003). One might also reduce unexplained variance in political preferences and behavior by identifying individual differences in trade-offs between self-interest and different moral preferences. In incentivized experiments, Fowler and Kam (2007) find that political participation is predicted by individual willingness to deviate from self-interest and act on altruistic preferences, and Dawes and Fowler (2007) establish further patterns in political participation based on individual differences in utilitarian versus Rawlsian moral preferences. Adding spectator treatments might help better identify moral preferences and explain stakeholder behavior, e.g., in Konow (2000), spectator decisions provided more precise measures of moral preferences and, in conjunction with stakeholder treatments, identified not only the impact of self-interest but also decomposed it into unadulterated (i.e., acknowledged) self-interest and self-deception (i.e., self-serving beliefs about what is right). One could similarly conduct experiments and surveys using spectators to disentangle and, thereby, better explain and predict the sources and effects of moral preferences on public opinion and political behavior.

There are various opportunities, in my mind, for empirical research on deliberative democracy and the quasi-spectator to enhance one another. For example, the quasi-spectator method offers a standard for calibrating the degree of impartiality present in the processes and outcomes of deliberation. As previously discussed, critics of deliberative democracy often question whether it will generate unbiased results, even if it produces consensus. Deliberative research typically adopts explicit steps to promote impartial discourse, e.g., by employing impartial moderators, providing balanced information and encouraging participants to remain open-minded. Indeed, some claim that balanced information is also a *product* of ideal deliberation. Be that as it may, the question remains whether there is an independent measure of

---

the current paper, "moral" refers not to what individuals prefer but rather to what they think *should* be, and not only with respect to questions like abortion but also on issues like the scale of government.

impartiality that could be employed to corroborate and perhaps improve the desired balance. Spectatorship offers such a standard, for example, individuals from outside a local community might be brought into a conflict over municipal planning, such as that described in Karpowitz and Mansbridge (2005), to hear the arguments of the stakeholders and to provide input on a fair resolution and, simultaneously, to help identify bias in different proposals. This application is to *outcomes*, but informed third parties could also be used to operate on the deliberative *process*. For instance, experiments on deliberation typically seek to provide balanced information, and informed third parties might help identify what information is genuinely balanced and impartial.

One outcome often expected from deliberation is opinion change, and this also represents an important area of overlapping research between deliberative democracy and public opinion. Nevertheless, deliberation sometimes does not produce significant opinion change and, in any case, critics question whether such change has normative value (e.g., if it is coerced). One possibility for addressing these questions is to engage third parties, perhaps parallel to stakeholders, in deliberation and gauge their opinions before and after. Since they are not stakeholders, any transformation in their views should not be due to coercion. In addition, such comparisons might also help clarify whether any absence of opinion change is due to a failure in stakeholder deliberation or, perhaps, to the fact that initial opinions were already unbiased.

Empirical deliberation research can also inform and strengthen quasi-spectator studies, particularly, I think, with respect to the epistemic aspects of impartiality. Stakes are often directly correlated with information about many factors that one expects to be relevant to arriving at impartial decisions. Stakeholders presumably know better what issues are most important and what arguments most salient to them. Sometimes they also possess more facts about the workings of the relevant people and institutions and how they might be affected by policies under consideration. Moreover, one of the virtues of deliberation heralded by its advocates is that the process itself helps to produce more thoughtful and informed views. One might, therefore, elicit the views of third parties who observe, and perhaps even participate in, deliberations with stakeholders in order better to understand and appreciate contextual factors.

Third parties might, in turn, be employed for epistemic purposes to improve deliberation. The results reported in this study indicate that sorting information into its relevant and irrelevant components using quasi-spectators offers opportunities for significant improvements in terms of unbiasedness and consensus among stakeholders. One proposal to reduce bias and dispute rates is to utilize quasi-spectators to identify the relevance of arguments advanced by one group of stakeholders and then to expose another group of stakeholders, who engage in deliberation, only to the relevant arguments.

The empirical methods inspired by spectator theory and deliberation theory have different relative benefits. It seems advisable, therefore, to make the choice of method, or combination of methods, reflect these context- and goal-dependent trade-offs. I argued here, for example, that the goal of spectator theory relates more to moral epistemology, including the search for the principles of ethics. Implementations of deliberation theory, on the other hand, have typically emphasized more contextually rich policy questions. The advantages of eliciting the views of disinterested parties are more salient in the former case, whereas the importance of parties being vested in understanding and grappling with complex issues is apparent in the latter case. This diversity of goals and contexts justifies, I believe, a variety of empirical methods, and a possible area for future work is to explore how different combinations of the quasi-spectator method, public opinion research and deliberation research might strengthen work on empirically informed ethics, public policy, policy preferences and voting.

REFERENCES

Amiel, Yoram, Frank A. Cowell, and Wulf Gaertner. 2008. "To Be or not to Be Involved: A Questionnaire-Experimental View on Harsanyi's Utilitarian Ethics," forthcoming in *Social Choice and Welfare*.

Ansolabehere, Stephen, Jonathan Rodden, and James M. Snyder. 2006. "Purple America." *Journal of Economic Perspectives* 20 (Spring): 97-118.

Ansolabehere, Stephen, Jonathan Rodden, and James M. Snyder. 2008. "The Strength of Issues: Using Multiple Measures to Gauge Preference Stability, Ideological Constrain, and Issue Voting." *American Political Science Review* 102 (May): 215-32.

Ashraf, Nava, Colin Camerer, and George Loewenstein. 2005. "Adam Smith, Behavioral Economist." *Journal of Economic Perspectives* 11 (Summer): 109-26.

Babcock, Linda, and George Loewenstein. 1997. "Explaining Bargaining Impasse: The Role of Self-Serving Biases." *Journal of Economic Perspectives* 11 (Winter): 109-26.

Barabas, Jason. 2004. "How Deliberation Affects Policy Opinions." *American Political Science Review* 98 (November): 687-701.

Brown, Vivienne. Forthcoming. "Agency and Discourse: Revisiting the Adam Smith Problem." In ed. Jeffrey T. Young. *Elgar Companion to Adam Smith*.

Dryzek, John S., and Simon Niemeyer. 2006. "Reconciling Pluralism and Consensus as Political Ideals." *American Journal of Political Science* 50 (July): 634-49.

Dunning, David, Judith A. Meyerowitz, and Amy D. Holzberg. 1989. "Ambiguity and Self-Evaluation: The Role of Idiosyncratic Trait Definitions in Self-Serving Assessments of Ability." *Journal of Personality and Social Psychology* 57 (December): 1082-90.

Faravelli, Marco. 2007. "How Context Matters: A Survey Based Experiment on Distributive Justice." *Journal of Public Economics* 91 (August): 1399-422.

Fishkin, James A., and Robert C. Luskin. 2005. "Experimenting with a Democratic Ideal: Deliberative Polling and Public Opinion." *Acta Politica* 40: 284-98.

Fishkin, James. 1991. *Democracy and Deliberation: New Directions for Democratic Reform*. New

Haven: Yale University Press.

Fleischacker, Samuel. 1999. *A Third Concept of Liberty: Judgment and Freedom in Kant and Adam Smith*. Princeton: Princeton University Press.

Fricke, Christel, and Hans-Peter Schutt, ed. 2005. *Adam Smith Als Moralphilosoph*. Berlin, New York: Walter de Gruyter, Inc.

Frohlich, Norman, and Joe A. Oppenheimer. 1992. *Choosing Justice: An Experimental Approach to Ethical Theory*. Berkeley, Los Angeles: University of California Press.

Frohlich, Norman, Joe A. Oppenheimer, and Cheryl L. Eavey. 1987. "Choices of Principles of Distributive Justice in Experimental Groups." *American Journal of Political Science* 31 (August): 606-36.

Fung, Archon. 2007. "Democratic Theory and Political Science: A Pragmatic Method of Constructive Engagement." *American Political Science Review* 101 (August): 443-58.

Goldstein, William M., and Elke U. Weber. 1995. "Content and Discontent: Indication and Implications of Domain Specificity in Preferential Decision Making." *Psychology of Learning and Motivation* 32: 83-136.

Griswold, Charles L. 1999. *Adam Smith and the Virtues of Enlightenment*. Cambridge: Cambridge University Press.

Haakonssen, Knud. 2006. *The Cambridge Companion to Adam Smith*. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, Sao Paulo: Cambridge University Press.

Habermas, Jurgen. 1990[1983]. *Moral Consciousness and Communicative Action*. Trans. Christian Lenhardt and Shierry Weber Nicholson. Cambridge: MIT Press.

Hanley, Ryan Patrick. 2008. "Enlightened Nation Building; the 'Science of the Legislator' in Adam Smith and Rousseau." *American Journal of Political Science* 52 (April): 219-34.

Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler. 1986. "Fairness as a Constraint on Profit Seeking: Entitlements in the Market." *American Economic Review* 76 (September): 728-41.

Karpowitz, Christopher F., and Jane Mansbridge. 2005. "Disagreement and Consensus: The Importance of Dynamic Updating in Public Deliberation." In ed. John Gastil and Peter Levine. *The Deliberative*

Democracy Handbook, 237-53. San Francisco: Jossey-Bass, Inc.

Knobe, Joshua. 2007. "Experimental Philosophy and Philosophical Significance." *Philosophical Explorations* 10 (June): 119-21.

Konow, James. 2000. "Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions." *American Economic Review* 90 (September): 1072-92.

Konow, James. 2003. "Which is the Fairest One of All?: A Positive Analysis of Justice Theories." *Journal of Economic Literature* 41 (December): 1188-239.

Konow, James. 2008. "Is Fairness in the Eye of the Beholder? An Impartial Spectator Analysis of Justice," forthcoming in *Social Choice and Welfare*.

List, Christian, Robert C. Luskin, James S. Fishkin, and Iain McLean. 2007. "Deliberation, Single-Peakedness, and the Possibility of Meaningful Democracy: Evidence from Deliberative Polls," Department of Government, London School of Economics, University of Texas at Austin, Stanford University, Oxford University, 4/January.

Mansbridge, Jane J. 1980. *Beyond Adversary Democracy*. New York: Basic Books, Inc.

Michelbach, Philip A., John T. Scott, Richard E. Matland, and Brian H. Bornstein. 2003. "Doing Rawls Justice: An Experimental Study of Income Distribution Norms." *American Journal of Political Science* 47 (July): 523-39.

Mitchell, Gregory, and Philip E. Tetlock. 2006. "Experimental Political Philosophy: Justice Judgments in the Hypothetical Society Paradigm," University of Virginia School of Law; Haas School of Business at University of California, Berkeley, 29/June.

Mouffe, Chantal. 1996. "Democracy, Power and 'the Political'." In ed. Seyla Benhabib. *Democracy and Difference: Contesting the Boundaries of the Political*, 245-56. Princeton: Princeton University Press.

Nadelhoffer, Thomas, and Eddy Nahmias. 2007. "The Past and Future of Experimental Philosophy." *Philosophical Explorations* 10 (June): 123-49.

Neblo, Michael A. 2007. "Change for the Better? Linking the Mechanisms of Deliberative Opinion Change to Normative Theory," http://polisci.osu.edu/faculty/mneblo/papers/ChangeC4.pdf.

Nichols, Shaun. 2004. "After Objectivity: An Empirical Study of Moral Judgment." *Philosophical*

Psychology 17 (March): 5-28.

Nyborg, Karine, and Kjell Arne Brekke. 2007. "Selfish Bakers, Caring Nurses? A Model of Moral Motivation," manuscript, University of Oslo, Norway.

Parrish, John M. 2007. *Paradoxes of Political Ethics: From Dirty Hands to the Invisible Hand*. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, Sao Paulo: Cambridge University Press.

Raphael, D. D. 2007. *The Impartial Spectator*. Oxford: Clarendon Press.

Rasmussen, Dennis C. 2006. "Does 'Bettering Our Condition' Really Make Us Better Off? Adam Smith on Progress and Happiness." *American Journal of Political Science* 100 (August): 309-318.

Rawls, John. 1971. *A Theory of Justice*. Cambridge: Belknap Press of Harvard University Press.

Rawls, John. 2000. *Lectures on the History of Moral Philosophy*. Cambridge, London: Harvard University Press.

Shapiro, Ian. 2002. "Problems, Methods, and Theories in the Study of Politics, or What's Wrong with Political Science and What to Do About It." *Political Theory* 30 (August): 596-619.

Smith, Adam. 1759 [1809]. *The Theory of Moral Sentiments*. Glasgow: R. Chapman.

Sulkin, Tracy, and Adam F. Simon. 2001. "Habermas in the Lab: A Study of Deliberation in an Experimental Setting." *Political Psychology* 22 (December): 809-26.

Traub, Stefan, Christian Seidl, Ulrich Schmidt, and Maria Vittoria Levati. 2005. "Friedman, Harsanyi, Rawls, Boulding - or Somebody Else? an Experimental Investigation of Distributive Justice." *Social Choice and Welfare* 24: 283-309.

Young, Iris. 1996. "Communication and the Other: Beyond Deliberative Democracy." In ed. Seyla Benhabib. *Democracy and Difference: Contesting the Boundaries of the Political*, 120-37. Princeton: Princeton University Press.

Zaller, John, and Stanley Feldman. 1992. "A Simple Theory of the Survey Response: Answering Questions Versus Revealing Preferences." *American Journal of Political Science* 36 (August): 579-616.

Appendices

Composite Questions

Key to Composite Questions
The text assigning allocator roles appears in **[**brackets**]** with **H** for High Stakeholder, **L** for Low Stakeholder and **S** for Spectator (where no **S** appears, there is no additional text for this role and the Spectator is by default). Text related to information conditions appears in **{**braces**}**, where **R** denotes Relevant information, **I** denotes Irrelevant information, and the remaining text is the Base information that is found in all versions.

Instructions

This questionnaire consists of several questions each describing a different scenario. Please read each question carefully, and then supply a numerical answer in the space provided. Please give exactly one answer to every question, as we cannot use forms with multiple or incomplete answers. This is not a test of knowledge or ability. Instead, we are interested in what you think should be done in each scenario given the information provided.

After you complete the questions, there is a final page requesting subject information. When you are finished, please put your form and pencil down and wait quietly. When everyone is finished you will individually and confidentially deposit your forms in the box in the front.

1. The Environmental Protection Agency (or EPA) is responsible for regulating the discharge of degradable waste by a pulp mill into a river. **[S:** The pulp mill involved is located in a different region of the country. / **H:** You live in a house, built by your great-grandfather, that is located near this river downstream from the pulp mill. / **L:** Suppose you and others commute from a neighboring county to work at this pulp mill, since there is no other industry or major employer in this region. You are the sole provider for yourself, your infant children and your elderly parents.**]** The EPA must decide whether to require the pulp mill to reduce its waste discharges into the river and, if so, by how much. Doing so would reduce various adverse effects of the discharge, but complying with EPA requirements would also require the pulp mill to cut its labor force of 400 workers and, perhaps, to close down altogether. **{R:** Cutting the waste by 30% would eliminate the noxious odors coming from the river but would result in the unemployment of 10 workers at the pulp mill. Cutting the waste by 60% would also make the river safe for drinking, swimming and fishing, but would cause a total of 20 workers to be laid off. Eliminating the waste altogether (that is, reducing it by 100%) would allow the return of an additional type of fish valued by some sports fishermen but would make the pulp mill unprofitable so that it would have to close down and lay off all 400 of its workers.**} {I:** The local labor union opposes any regulation of the pulp mill. The local homeowners' association is campaigning to have the discharges eliminated altogether.**}** By how much, if any, do you think the EPA should require the pulp mill to *reduce* its discharges (Enter a number from 0% for "no reduction" to 100% for "complete elimination" in the space below)?

_____ %

2. **[S:** You are the judge deciding the outcome of a civil suit brought by a motorcyclist against the driver of a car that hit him. / **H:** You have brought a civil suit against the driver of a car that hit you while you were on your motorcycle. / **L:** While driving your car, you hit a motorcycle, and that motorcyclist has brought a civil suit against you.**]** The suit demands $100,000 in damages for medical expenses, loss of earnings and pain and suffering (vehicle repairs were covered by insurance), but the actual award could be anything between $0 and $100,000. In court testimony, the facts have been presented as follows. The motorcyclist pulled out of a parking lot into a street a few feet from a stop sign and was thrown from his motorcycle when the car struck him. **{R:** As a result of the accident, the motorcyclist has lost earnings of about $3,000 due to missed work time and has incurred medical expenses of around $12,000.**} {I:** A

passenger in the car claims the motorcycle pulled in front of the car, but a bystander denies this and says that the car was speeding.**}** How much do you think the court should require **[S,H:** driver of the car / **L:** you**]** to pay the **[S,L:** motorcyclist / **H:** you**]** (Enter a number from $0 to $100,000)?

$ _____

3. **[H:** Suppose that you rely exclusively on the local hospital's preventative services since your job does not provide health insurance or enough income for you to buy insurance on your own. These preventative services include screening that enabled you to treat tuberculosis before it became life threatening and, for your child, immunizations and medication to control her diabetes. / **L:** Suppose you frequently rely on the local hospital's emergency room when your child experiences life-threatening asthma attacks.**]** A hospital budget committee must decide how much of the budget it controls to allocate to the hospital's emergency services versus to its preventive services for the community. **{R:** At present, many patients in the community go to the emergency room for their non-emergency needs because they are uninsured. By increasing the budget to preventative services to 60%, the needs of these patients would be covered, and the reduced burden on emergency services would allow it to provide almost the same level of services as previously.**} {I:** Currently, patients seeking preventative services must schedule appointments six months or more in advance.**}** What percentage of the budget do you think should be allocated to *preventative services* (Enter a number from 0% to 100% in the space below)?

_____ %

4. Suppose **[S:** Adam and Bill / **H:** you and Bill / **L:** you and Adam**]** worked last weekend stuffing envelopes for a mass mailing. This job took a total of 11 man hours, but **[S,L:** Adam / **H:** you**]** worked more hours than **[S,H:** Bill / **L:** you**]**. **{R:** Specifically, **[H:** you / **S,L:** Adam**]** worked 8 hours whereas **[S,H:** Bill / **L:** you**]** worked 3 hours.**} {I:** To get to the job, **[H:** you / **S,L:** Adam**]** drove two hours and **[S,H:** Bill / **L:** you**]** took a one hour bus ride.**}** The total pay for this 11 hour job is $100. How much of this $100 do you think **[S:** Adam and Bill / **H:** you and Bill / **L:** you and Adam**]** should each receive (Enter amounts for each person below and make sure the two amounts total $100)?

[S: Adam / **H,L:** You**]**	$ S H
[S,H: Bill / **L:** Adam**]**	$ L
Total	$100

5. [**H:** Suppose that you recently lost both parents but that, with the help of financial assistance you have received from the state, you just finished school at the top of your class. / **L:** Suppose that, having worked your way through school, you are now work long hours at a challenging but modestly paid job at a bank.**]** The state provides support to those in need for a limited period of time. For example, John, who needs one year to complete a high school diploma, is eligible to receive such support. **{R:** The state has determined that the basic needs of a person living in this area for food, housing and clothing equal $800 per month.**} {I:** Because of a downturn in the economy, tax revenues have fallen and the number of people applying for support from the state has increased.**}** How much do you think the state should provide in total support for John per month (Enter a number from $0 to $1000)?

$ _____ per month

6. A large company has two divisions. The one division produces film for traditional cameras, which is the business the company was founded on. The other, newer division is focused on technologies for digital photography and printing. **[H:** Suppose your only employable skills are in the film division, where you have worked for 25 years. You are eligible for retirement benefits in 3 years, which you will lose if the film division is scaled back and you are laid off. / **L:** Suppose that when you saw the move toward digital technologies, you went back to college at your own cost to acquire computer skills and are now employed in the digital division.**]** Due to changing consumer demand, the traditional film division is on the decline and its share of company revenues is falling. The company's budget for plant, machinery and equipment in the coming year totals $10 billion, and its board must decide how much of this to devote to

the film division and how much to the digital division. **{R:** Company finance analysts expect revenues from the film division to fall from 60% currently to only 10% in five years. In order to protect the company's financial health and survival, they recommend focusing expenditures for plant, machinery and equipment on the digital division and devoting $9 billion of next year's budget to the digital division and only $1 billion to the film division.**} {I:** Efficiency in the film division could be improved slightly by minor expenditures on tools, but its plant, machinery and equipment are otherwise up-to-date.**}** How much of this $10 billion do you think the board should budget for the *film* division of the company (Enter a number in billions of dollars from 0 to 10)?

$ _____ billion

Subject information

Please answer all questions, indicating just one answer per question, as we cannot use forms with incomplete or multiple answers.

1. What is your college?

1 Business
2 Communications and Fine Arts
3 Liberal Arts
4 Science and Engineering

2. What is your first major (if undeclared, write UD)?

_____

3. What year in college are you?

1 Freshman
2 Sophomore
3 Junior
4 Senior
5 Graduate

4. What is your age?

_____ years

5. What is your gender?

1 Male
2 Female

6. What is your ethnicity (if several apply, please choose the <u>one</u> that you consider most accurate)?

1 Asian/Pacific-Islander
2 Black/African-American
3 Caucasian
4 Latino/Hispanic
5 Middle-Eastern
6 Native-American/American Indian

7. What is your best estimate of your total expenditures this school year (September through May)? Please consider all expenses including tuition, housing, food, clothing, transportation, entertainment, etc., even if some are covered by financial aid or grants.

$_____ for the current school year (September through May)

8. What is the total (gross) income last year of your parents or guardians (or spouse, if married)? Exclude your own earnings. Please choose a <u>single</u> response, even if it is a guess.

1 $0 to less than $25,000
2 $25,000 to less than $50,000
3 $50,000 to less than $75,000
4 $75,000 to less than $100,000
5 $100,000 to less than $125,000
6 $125,000 to less than $150,000
7 $150,000 or more

9. How many hours per week do you usually work (Enter 0 if none)?

_____ hours per week

10. Approximately how much money have you earned total through your work over the past year (the past twelve months)?

$ _____

TABLE A1. Demographic Characteristics of the Sample and the Population

Characteristic	Sample	Population
Ethnicity	%	%
Asian/Pacific-Islander	11	11
Black/African-American	7	7
Caucasian	61	54
Latino/Hispanic	18	18
Other/Decline to State	3	10
Gender		
Female	63	60
Male	37	40
College		
Liberal Arts	38	37
Communications and Fine Arts	17	25
Business Administration	35	23
Science and Engineering	10	15

Notes: The sample consists of students from general classes in psychology and economics at a comprehensive US university that has an annual undergraduate enrollment of about 5700. The population percentages are for the academic year in which the largest number of observations were collected.

TABLE A2. Relevant and Irrelevant Information: Tests of Differences in Means

	H_0: S = SR		H_0: SR = SRI	
Question	t-statistic	p-value	t-statistic	p-value
1	−6.23	.000	1.23	.222
2	5.96	.000	−0.83	.409
3	−6.86	.000	0.00	.997
4	−15.43	.000	1.07	.288
5	−9.89	.000	1.48	.142
6	5.45	.000	−0.40	.691

Note: P-values are based on two-tail tests.

TABLE A3. Regression Results for Figures 2 and 3

	(1) Mean H – Mean L Response (Figure 2)	(2) Variance (Figure 3)
Question 2	10.48**	213.1*
Question 3	−4.60	−334.6**
Question 4	−3.80	−468.3**
Question 5	7.10*	56.9
Question 6	−2.65	−215.3*
Constant	1.53	709.9**
Base Information	11.97**	
Irrelevant	5.75*	
Relevant + Irrevelant	6.27*	
Stakeholder R		−290.6**
Stakeholder I		−73.0
Stakeholder R+I		−280.4**
Spectator		−48.1
Spectator R		−345.2**
Spectator I		−47.8
Spectator R+I		−284.7**
Observations	24	48
Adjusted R-squared	0.71	0.70

Notes: ^ $p<0.10$, * $p<0.05$, ** $p<0.01$; the constant in regression (1) is the estimated coefficient for the omitted category R and in regression (2) is the estimated coefficient for the omitted category of Stakeholder variance under the Base information condition.

TABLE A4. Regression Analysis: Moral Judgments of Spectators

	Question					
Regressors	1. Environ.	2. Law	3. Bioethics	4. Accountab.	5. Need	6. Efficiency
Relevant	18.96**	−21.71**	14.14**	13.37**	32.95**	−12.07**
Relev+Irrelev	16.63**	−17.38**	14.18**	12.23**	27.87**	−10.51**
Irrelevant	6.97^	0.62	−2.21	−1.30	9.31*	−8.13**
Gender	4.55^	2.12	2.21	−0.40	−1.02	3.02
Nonwhite	−3.64	0.61	1.27	−1.70*	8.63**	1.40
Business	0.63	1.56	−1.42	−1.52^	−4.69	−2.65
Comm/FineArts	0.56	−0.30	−1.47	−0.64	0.31	1.78
Science/Engin	−1.50	0.21	−1.78	0.50	−1.76	5.48
Class	−1.34	0.97	0.89	−0.62	−2.10	−1.41
Age	0.34	1.23	0.64	0.13	1.97*	−1.24
Expen($1000/yr)	−0.13	0.15	0.04	−0.00	−0.02	−0.06
Parent income	−0.62	−0.36	−0.67	−0.21	0.51	0.56
Hours work/wk	−0.13	−0.06	0.13	0.02	0.06	0.06
Earn($1000/yr)	−0.14	−0.14	−0.08	−0.04	−0.18	0.23
Observations	309	332	314	327	310	324
R-squared	0.15	0.13	0.25	0.59	0.31	0.13

Notes: The entries are regression coefficients, whereby omitted categories for the dummy variables are white, male and Liberal Arts College.
^p<.10, *p<.05, **p<.01 (two-tailed)

TABLE A5. Regression Analysis: Moral Judgments of High Stakeholders

	Question					
Regressors	1. Environ.	2. Law	3. Bioethics	4. Accountab.	5. Need	6. Efficiency
Relevant	10.37*	−35.59**	9.97**	10.64**	14.14**	−14.91**
Relev+Irrelev	12.10**	−27.00**	15.69**	9.60**	7.82*	−15.33**
Irrelevant	0.17	12.69*	−0.52	0.85	−6.15	−8.00*
Gender	1.66	−2.55	0.79	−0.96	−1.02	6.54*
Nonwhite	−6.91*	2.68	4.14	−0.44	−0.47	2.46
Business	3.30	0.82	1.96	−2.01	−1.07	−3.49
Comm/FineArts	11.48*	−4.17	3.98	1.20	4.63	5.09
Science/Engin	0.15	−3.71	−0.98	0.82	−2.73	−6.50
Class	3.54	−10.65*	0.79	−0.24	5.15^	2.37
Age	−2.44	8.08*	−1.64	−0.05	−4.43*	−0.34
Expen($1000/yr)	−0.17	0.08	−0.02	−0.02	−0.12	−0.12
Parent income	1.13	0.52	−0.97	0.00	0.09	0.76
Hours work/wk	−0.12	0.34	0.05	0.02	0.09	0.27^
Earn($1000/yr)	−0.61	0.12	−0.08	0.12	−0.86	−0.61^
Observations	225	219	226	223	221	223
R-squared	0.13	0.12	0.22	0.33	0.16	0.24

Notes: The entries are regression coefficients, whereby omitted categories for the dummy variables are white, male and Liberal Arts College.
^p<.10, *p<.05, **p<.01 (two-tailed)

TABLE A6. Regression Analysis: Moral Judgments of Low Stakeholders

	Question					
Regressors	1. Environ.	2. Law	3. Bioethics	4. Accountab.	5. Need	6. Efficiency
Relevant	19.34**	−15.43**	16.71**	10.27**	32.40**	1.06
Relev+Irrelev	16.71**	−15.88**	13.53**	9.50**	23.06**	−11.68**
Irrelevant	5.87	16.08**	−0.01	−0.35	8.73^	−0.02
Gender	0.89	5.43	−1.38	1.42	4.23	1.81
Nonwhite	−3.28	0.98	−1.23	−2.35	8.25*	−1.59
Business	5.99	−2.78	−3.15	−2.01	−3.48	1.20
Comm/FineArts	4.68	−3.69	−0.39	−0.42	−4.32	6.00^
Science/Engin	−0.03	4.12	0.48	−3.66	−8.52	−8.76^
Class	2.69	−4.88	−0.72	4.77*	−6.78	2.85
Age	−2.04	7.12^	0.58	−2.62^	3.13	−0.05
Expen($1000/yr)	−0.05	0.13	−0.01	−0.04	0.01	−0.01
Parent income	−0.73	0.32	−0.69	−0.01	−0.54	0.32
Hours work/wk	−0.04	−0.15	−0.11	0.09	−0.20	0.30*
Earn($1000/yr)	0.43	0.32	−0.08	−0.19	0.21	−0.34
Observations	224	216	238	220	215	227
R-squared	0.11	0.28	0.25	0.21	0.26	0.14

Notes: The entries are regression coefficients, whereby omitted categories for the dummy variables are white, male and Liberal Arts College.
^p<.10, *p<.05, **p<.01 (two-tailed)

TABLE A7. Tests of Differences in Estimated Coefficients for Regression (2) in Table A3

	Relevant Stakeholder	Irrelevant Stakeholder	Relev+Irrel Stakeholder	Base Spectator	Relevant Spectator	Irrelevant Spectator	Relev+Irrel Spectator
Base Stakeholder	−2.95 0.0060	−0.74 0.4640	−2.85 0.0070	−0.49 0.6280	−3.50 0.0010	−0.49 0.6300	−2.89 0.0070
Relevant Stakeholder		4.87 0.0339	0.01 0.9188	6.06 0.0189	0.31 0.5827	6.07 0.0188	0.00 0.9525
Irrelevant Stakeholder			4.43 0.0425	0.06 0.8017	7.63 0.0091	0.07 0.7996	4.61 0.0387
Relev+Irrel Stakeholder				5.56 0.0241	0.43 0.5153	5.57 0.0239	0.00 0.9662
Base Spectator					9.09 0.0048	0.00 0.9978	5.76 0.0218
Relevant Spectator						9.11 0.0047	0.38 0.5428
Irrelevant Spectator							5.78 0.0217

Notes: The top entries in the first (Base Stakeholder) row are t-statistics and F-statistics otherwise; the bottom entries are p-values.