

The Dark Side of Morality: Group Polarization and Moral Epistemology

Marcus Arvan
University of Tampa
marvan@ut.edu

Abstract: This article argues that philosophers and laypeople commonly conceptualize moral truths or justified moral beliefs as discoverable through intuition, argument, or some other purely cognitive or affective process. It then contends that three empirically well-supported theories all predict that this ‘Discovery Model’ of morality plays a substantial role in causing social polarization. The same three theories are then used to argue that an alternative ‘Negotiation Model’ of morality—according to which moral truths are not discovered but instead *created* by actively negotiating compromises—promises to reduce polarization by fostering a progressive willingness to ‘work across the aisle’ to settle moral issues cooperatively. This article then examines potential methods for normatively evaluating polarization, arguing there are *prima facie* reasons to favor the Negotiation Model over the Discovery Model based on their hypothesized effects on polarization. Finally, I outline avenues for further empirical and philosophical research.

Most of us are accustomed to thinking of morality in a positive light. Morality, we say, is a matter of distinguishing good from bad, right from wrong, and acting well. However, moral beliefs and discourse also plausibly play a role in *group polarization*, the tendency of social groups to divide into progressively more extreme factions, each of which regards other groups to be ‘wrong.’¹ Group polarization often occurs along moral lines², and is known to have many disturbing social effects, including increasing racial prejudice among the already moderately prejudiced³, group decisions that are more selfish, competitive, and less trusting than individual decisions⁴, the erosion of public trust⁵, juries imposing more severe

¹ See Isenberg (1986) and Pruitt (1971).

² See Haidt (2001, 2012).

³ Myers and Bishop (1970).

⁴ Luhan, Kocher, and Sutter (2009).

⁵ Rapp (2016).

punishments in trials⁶, extreme political decisions⁷, and war, genocide, and other violent behavior.⁸

This article argues that three empirically-supported theories of group polarization all predict that polarization is likely caused in significant part by the widespread acceptance of a *Discovery Model* of morality: a model which holds that moral truths or morally justified beliefs can be discovered through moral intuition, moral reasoning, or some other purely cognitive or affective process. §1 of this paper clarifies this model, showing how it appears commonly presupposed by laypeople and in moral philosophy, cohering as well with empirical research on how people ordinarily form moral beliefs. §2 then argues that three leading empirical theories of group polarization—Social Comparison Theory⁹, Informational Influence Theory¹⁰, and Self-Categorization Theory¹¹—all predict that the Discovery Model likely plays a significant role in causing polarization. Next, §3 argues there are converse reasons to believe that an alternative *Negotiation Model* of morality, according to which moral truths are instead created by negotiating compromises, would likely prevent polarization and promote cooperative compromise. §4 then outlines potential methods for normatively evaluating polarization and the two models of morality, suggesting that the Negotiation Model's theorized effects on polarization are *prima facie* desirable and the Discovery Model's *prima facie* undesirable. Finally, §5 explores avenues for further empirical and philosophical research.

⁶ Bray and Noble (1978)

⁷ Walker and Main (1973)

⁸ Newman (2002), Sunstein (2002).

⁹ Bray and Noble (1978), Mackie (1986).

¹⁰ Stoner (1961); Myers and Arenson (1972); Hinsz and Davis (1984).

¹¹ Abrams et al (1990).

As a point of clarification, this paper's aims are modest.¹² First, it neither aims to establish that the Discovery Model causes polarization, nor that the Negotiation Model is psychologically realistic or would reduce polarization. Because these are complex empirical issues, this article merely aims to provide strong theoretical grounds for investigating these matters further in future research. This paper also does not aim to settle a variety of philosophical questions, including whether or when polarization is morally desirable, or whether the Negotiation Model should be favored over the Discovery Model *simpliciter*. Although §4 contends there are *prima facie* reasons to favor the Negotiation Model over the Discovery Model due to their theorized effects on polarization, these hypotheses must also be tested in future research. Finally, although some readers may worry that this article engages in *hineinterpretieren*—a problematic form of *post-hoc* theorizing that involves drawing spurious connections between theories and data—this article is better understood as embodying a common and important method: that of arguing that genuine conceptual connections between existing theories and data entail novel, testable empirical hypotheses—a standard first step in justifying new empirical research programs.¹³

1 The Discovery Model of Morality

Some philosophers¹⁴ and everyday laypeople purport to be moral skeptics, alleging that there are no moral facts at all. Nevertheless, most people appear to form and express moral beliefs in the course of everyday life. We regularly speak of people doing 'right', 'wrong', 'good', and 'bad.' We also tend to do so in accordance with a particular model of morality: a Discovery Model according to which moral truths or morally justified beliefs *exist to be*

¹² I thank two anonymous reviewers for inviting me to rethink and foreground my argument's scope.

¹³ See e.g. Einstein (2005) and Higgs (1964) for two particularly famous cases.

¹⁴ See e.g. Joyce (2007, 2015) and Mackie (1977).

discovered through intuition, moral reasoning, or some other cognitive or affective process. The Discovery Model, as I propose we understand it, does not hold that we come to believe moral propositions passively or unreflectively. It is instead the conjunction of the following two claims:

The discovery model of meta-ethics: there are in general *truths* or *justified beliefs* about moral issues (e.g. about right, wrong, good, bad) that can in principle be ascertained *unilaterally*, via an individual's own intuition, use of philosophical argumentation (including normative ethical theorizing), or some other cognitive or affective process(es).

The discovery model of moral-belief formation: people who tacitly or explicitly endorse the discovery model of meta-ethics will tend to *form moral beliefs* through intuition, argument, or other such "unilateral discovery process", and believe at a higher-order level that they have *discovered moral truths* or developed *justified moral beliefs* through some such process.

We can see just how commonly the Discovery Model is presupposed by examining everyday life, academic moral philosophy, and social-psychological research.

Consider first everyday moral practice. When it comes to applied moral issues, laypeople and philosophers typically presuppose that there are 'moral facts'—either moral truths or facts about which moral beliefs are justified—that can be discovered through intuition, argument, or some other cognitive or affective process. For example, laypeople who believe that abortion is morally wrong often claim to base this 'discovery' on the intuition or some argument that human beings have a *moral right to life*, arguing that since

fetuses are human beings, we should all believe that abortion is wrong.¹⁵ Conversely, people who believe abortion is morally permissible typically appear to base their opposing belief on different intuitive or argumentative ‘discoveries’, such as that fetuses do not have a moral right to life at certain stages of development¹⁶, or alternatively, that a fetal right to life is not a right to depend on a mother’s body.¹⁷ To take another case, consider the Israeli-Palestinian conflict. Here too, both sides appear to treat moral truths as discoverable through intuition or argument. Whereas pro-Palestinian voices defend the moral ‘discovery’ Israel has wrongly occupied Palestinian lands, pro-Israelis typically defend the opposite ‘discovery’: that Israelis have *rightly* occupied Palestine, claiming that settlers are simply ‘living on land that Israel has liberated.’¹⁸ In each case, we see the Discovery Model’s two components exemplified: its meta-ethical component that moral truths or morally justified beliefs can be discovered through cognitive or affective process, as well as its psychological account of moral-belief formation (people forming moral beliefs representing their supposed ‘moral discoveries’). These are not isolated cases. The Discovery Model appears presupposed in the moral teachings of major world religions—with Judaism holding that we can discover moral truths through the Ten Commandments; Christianity that we can discover moral truths via Christ; Islam that moral truth is to be discovered through the Quran; Buddhism that moral truth is to be found in the Noble Eightfold Path; etc.

The Discovery Model also clearly pervades academic moral philosophy. For example, in contemporary meta-ethics, the notion that moral facts exist to be discovered is central to

¹⁵ See e.g. NRLC (2016), Pro-Life Perspective (2016).

¹⁶ See e.g. Arthur (2001).

¹⁷ See e.g. Liberty Women’s Health (2016)

¹⁸ See Black, Wedeman, and Mullen (2015) for a brief overview.

many theories of moral semantics and moral epistemology. While non-cognitivist interpretations of moral language exist, the dominant metaethical view of moral language is cognitivism: the view that sentences of the form, “X is morally wrong”, are true just in case it is a *fact* that X is morally wrong.¹⁹ Many moral realists argue that cognitivism is not only true, but that moral facts are mind-independent features of the world discoverable by us.²⁰ Further, consider other theories of moral epistemology: moral intuitionism, constitutivism, and reflective equilibrium. Intuitionists hold that moral truths can be discovered through moral perception²¹; constitutivists that moral truths can be discovered by reference to constitutive features of agency²²; and proponents of reflective equilibrium that we can discover which moral beliefs are justified through a process of reflecting back and forth on moral principles, moral judgments about cases, and theoretical considerations.²³

Now consider normative ethical theory. Here too the Discovery Model is commonly endorsed. Act-utilitarians hold that an action is right if and only if the act maximizes utility—facts that can in principle be discovered. Kantians hold that an action is permissible if and only if its maxim can be willed as a universal law²⁴, respecting the humanity of oneself and others²⁵—facts that can also in principle be discovered.²⁶ Aristotelian virtue ethicists hold moral virtues are beneficial character traits which we can *discover* to be necessary for living well.²⁷ And so on. Many other influential approaches to normative ethics—contractualism,

¹⁹ See van Roojen (2015) for an overview.

²⁰ See e.g. Shafer-Landau (2003), Brink (1989), Dancy (1986), Finlay (2007).

²¹ See e.g. Audi (2015) as well as Stratton-Lake (2014): introduction, for an overview.

²² See e.g. Kant (1785, 1797), Korsgaard (2008, 2009), and Katsafanas (2011).

²³ Daniels (2010). I thank an anonymous reviewer for encouraging me to highlight this.

²⁴ Kant (1785): 4:421.

²⁵ Ibid: 4:429.

²⁶ Ibid: 4:422-3, 4:429-31.

²⁷ See *Nicomachean Ethics*. Also see Hursthouse (1999): ch. 1.

Rossian pluralism, moral particularism, etc.—similarly hold that moral truths can be discovered by intuition, argument, or some other cognitive or affective process.²⁸

The Discovery Model also dominates theories of political morality—specifically, theories of justice. For example, Rawls argues we can discover through a process of rational reflection ('reflective equilibrium') that a just society would conform to two liberal-egalitarian principles of justice.²⁹ Robert Nozick argues we can discover through rational reflection that justice instead requires a minimal libertarian state.³⁰ And so on.

The Discovery Model also dominates applied ethics. In the applied ethical literature on abortion, some argue that abortion can be discovered to be wrong because it violates the moral rights of the fetus³¹; others argue that abortion can be discovered to be morally permissible at certain stages of fetal development due to the fetus not being a person³²; others still argue that abortion is permissible because a fetal right to life is not a right to depend on a mother's body.³³ Indeed, the applied ethics literature is replete with works arguing that we can discover truths about applied ethical issues—about the ethics of torture, gun control, warfare, treatment of animals, etc.—through intuition, argument, or other cognitive or affective processes.

Finally, the Discovery Model's account of moral-belief formation coheres with the dominant empirical model of moral-belief formation in social psychology: the Social Intuition Model (SIM) which holds that moral beliefs are the result of sudden flashes of

²⁸ See e.g. Scanlon (1998): 4, 191; Ross (1930); Dancy (2013); and Parfit (2011): §49.

²⁹ Rawls (1971).

³⁰ Nozick (1974).

³¹ See e.g. Pojman (1998) and Marquis (2007).

³² See e.g. Warren (1973).

³³ Thomson (1976).

affectively laden intuitions, with moral reasoning largely serving a subservient role to justify one's beliefs *ex post facto*.³⁴

Several points are important to clarify before examining in §2 how the Discovery Model is likely involved in polarization.³⁵ First, my claim is not that laypeople explicitly accept the Discovery Model as a philosophical theory. This is important to recognize, as it might seem implausible that an abstract philosophical theory that the average layperson has never heard of could produce group polarization among laypeople. My claim, to be clear, is merely that laypeople implicitly *conform* to the Discovery Model in their behavior—specifically, in the simple, common acts believing they have discovered moral truths or morally justified beliefs (e.g. 'It's not just my opinion, abortion is wrong!'), and in forming first-order moral beliefs (e.g. 'I believe that abortion is wrong'). The examples enumerated above—of how laypeople tend to think about moral issues ranging from abortion, to the Israeli-Palestinian conflict, to the moral teachings of major religions—illustrate just how common these Discovery-Model-conforming behaviors are among laypersons.

Second, the Discovery Model should not be conflated with moral realism—the view that moral truths exist objectively as 'mind-independent' facts or properties. To see why, consider the divine command theory of ethics, the view that God's will or commands define what is morally right, wrong, good, and evil. Insofar as divine command theory makes morality dependent on God's mind, it is clearly not a form of mind-independent moral realism. Nevertheless, it *does* conform to the Discovery Model, at least for human agents. For although divine command theory holds that God's will constitutes the normative basis of

³⁴ Haidt (2001).

³⁵ I thank a reviewer for encouraging me to clarify the issues that follow.

moral norms, divine command theory also holds these are norms *for humans to discover* through cognitive or affective processes—specifically, through discovering what God’s will or commands actually are. Indeed, ordinary laypeople who are committed to divine command theory are often clearly committed, at least implicitly in their behavior, to both components of the Discovery Model: its model of meta-ethics and its model of moral-belief formation. For example, one very common religious argument against the moral permissibility of abortion is that we can *discover* through religious scripture or tradition that abortion is wrong because it is against God’s will. For example, one common Christian argument³⁶ against abortion is that the Bible simultaneously implies that fetuses have souls (e.g. ‘Before I formed you in the womb I knew you’³⁷) while asserting God’s Seventh Commandment, ‘You shalt not kill’—two claims that, or so it is alleged, jointly entail that abortion is morally wrong *because* it is against God’s will.³⁸ According to those who subscribe to this argument, we can *discover* the moral fact abortion is wrong by reading the Bible (viz. the discovery model of meta-ethics). People who endorse this line of argument also appear to *form their first-order moral beliefs* (e.g. about abortion’s wrongfulness) on this basis of this ‘moral discovery’ (viz. the discovery model of moral-belief formation). Similar commitments to the Discovery Model are exhibited by scripture- and tradition-based arguments espoused by members of other major religions, including Judaism³⁹, Islam⁴⁰, Hinduism⁴¹, and so on.

Third, it is important to recognize that this article does not argue that the Discovery Model is the only cause of group polarization. There may be multiple sources of polarization,

³⁶ See e.g. Keseman (n.d), *Eternal Word* (1996), etc.

³⁷ *The New American Bible*, Jeremiah 1:5

³⁸ *Ibid*, Exodus 20:13.

³⁹ See e.g. BBC BiteSize (n.d.)

⁴⁰ BBC (2009a).

⁴¹ BBC (2009b)

including other sources that might even interact with or magnify the Discovery Model's hypothesized polarizing features. For example, if the present article is correct, then subscribing to divine command theory (e.g. as many religious believers appear to do) should generate polarization at least in part because people who subscribe to it conform to the Discovery Model. However, this source of polarization is consistent with and might interact with other polarizing forces, such as personal allegiance to God or belief that religious values are especially sacred.⁴² These features of religious belief or motivation might magnify the Discovery Model's polarizing features—by perhaps causing religious believers to become particularly unwilling to question their 'moral discoveries' (something that might make such individuals even more likely to adopt the polarizing behaviors described in §2 of this article). My argument, as such, is not meant to exclude other causes of polarization. It merely aims to establish that the Discovery Model is likely one significant cause of polarization—one that may complement or interact with other independent sources of polarization.

Finally, my argument does not hold that the Discovery Model generates polarization necessarily or as a matter of principle. Rather, my argument is that the Discovery Model *tends* to play a significant role in polarization due to contingent psychological facts about human beings. To see why this is important, consider an absolute Platonist about ethics, someone who believes that morality is a matter of discovering and conforming the Form of the Good. Notice that, at least in principle, one could be a moral Platonist of this sort and not be very confident in one's moral beliefs—for one might think (as Socrates appeared to in Plato's *Apology*⁴³) that it is very hard, perhaps even impossible, to know the Form of the Good.

⁴² See Sheikh et al. (2012).

⁴³ Plato, *Apology*: 21d.

Consequently, it seems possible that one might subscribe to the Discovery Model (as a ‘Humble Moral Platonist’) and yet not act in any of the polarizing ways described below in §2. However, although this is worth recognizing, it is also worth recognizing Humble Moral Platonists of this sort are not the norm, either among philosophers or among laypeople. On the contrary, as we see in the examples given earlier and in our daily lives, people typically appear to think that they *have* discovered moral truths or arrived at morally justified beliefs. For example, in public debate one rather tends to find people who claim to be very sure that (e.g.) abortion⁴⁴ and gun-control⁴⁵ are wrong, others who claim to be very sure of the opposite positions⁴⁶, etc. Similarly, moral philosophers who publish works on moral issues do not tend to be Humble Moral Platonists, either. Rather, they tend to defend *particular moral discoveries*—for instance, that there are mind-independent moral facts we can discover⁴⁷, that abortion is wrong⁴⁸, that women have a right to choose⁴⁹, etc.

My hypothesis, as §2 will present momentarily, is merely that these phenomena—that is, the way that people tend to actually conform to the Discovery Model, given contingent human psychological tendencies—serve to generate polarization. I am interested in this contingent hypothesis for several related reasons. First, polarization is not merely an abstract philosophical problem. It is a serious moral, social, and political issue at present, *in this world*—a world not widely populated by Humble Moral Platonists. Consequently, it is important to understand its actual sources. Second, there is ample empirical evidence that however epistemically humble human beings might be in principle, they tend not to be so

⁴⁴ Again see NRLC (2016), Pro-Life Perspective (2016).

⁴⁵ See e.g. Rowe (n.d.).

⁴⁶ See e.g. Feminist eZine (n.d.), Brezenoff (2018).

⁴⁷ See e.g. Shafer-Landau (2003), Brink (1989), Dancy (1986), Finlay (2007).

⁴⁸ Pojman (1998) and Marquis (2007).

⁴⁹ Thomson (1976).

epistemically humble given a wide variety of contingent psychological tendencies (see e.g. research on confirmation bias⁵⁰, belief conservation⁵¹, the overconfidence effect⁵², anchoring⁵³, and the bandwagon effect⁵⁴). My aim, as such, is merely to motivate a contingent hypothesis: that given what we know about common human psychological tendencies, the Discovery Model likely plays a very significant role in causing polarization. If correct, this is an important finding—as it may not only help us understand the actual (albeit contingent) causes of polarization, but also inform further inquiry into how to prevent it (one possibility being the Negotiation Model this paper explores, another possibility being changing human psychological tendencies, a third possibility being becoming Humble Moral Platonists, etc.).

2 The Discovery Model and Group Polarization

There are two leading empirical theories of the causes of group polarization: Social Comparison Theory and Informational Influence Theory.

According to Social Comparison Theory, individuals in groups have a psychological tendency to want to *gain acceptance* and be perceived favorably by other members of their group. This desire for acceptance causes individuals in the group to adopt ever-so-slightly more extreme views than those already typical in the group, so as to “impress” and “prove” themselves to other members of the group⁵⁵—something which often takes the form of *moral grandstanding*.⁵⁶ Group polarization then occurs when many individuals in the group do more or less the same thing, progressively adopting more extreme views to impress each

⁵⁰ Oswald & Grosjean (2004).

⁵¹ Ducharme (1970).

⁵² Hoffrage (2004).

⁵³ Zhang et al. (2007).

⁵⁴ Nadeau et al (1993).

⁵⁵ Bray and Noble (1978), Myers and Bishop (1970), and Luhan et al (2009).

⁵⁶ See Tosi & Warnke (2016).

other, causing the representative beliefs of the entire group to become progressively more extreme. Finally, and importantly, studies indicate that this phenomenon is even more likely to occur with respect to “judgmental issues”, such as moral or political matters.⁵⁷ For instance, a recent study on Twitter regarding the shooting of an abortion doctor indicated that like-minded individuals on both sides of the issue tend to group together, reinforcing and progressively polarizing pro-life and pro-choice views.⁵⁸

Informational Influence Theory supplements this account with a complementary mechanism: the tendency of people to group together with likeminded individuals to present novel *arguments and information* in favor of their preferred views, leading individuals in the group to become more easily convinced of even more extreme views.⁵⁹ For example, members of different political parties tend to frequent different news sources and social media networks.⁶⁰ These differential sources of information tend to provide members of each group with new information and arguments supporting their members’ preexisting views, ignoring or delegitimizing countervailing information and arguments, thereby making individuals in each groups more likely to develop more polarized views.⁶¹ Further, research indicates this mechanism is especially strong for “intellectant” issues—or issues involving intellectual debate, including moral issues.⁶² Informational Influence Theory also coheres with a well-established individual bias: *confirmation bias*, the tendency of people to

⁵⁷ Isenberg (1986).

⁵⁸ Yardi and Boyd (2010).

⁵⁹ See Vinokur and Burnstein (1974)

⁶⁰ See e.g. Iyengar & Hahn (2009).

⁶¹ See e.g. Morris (2007).

⁶² Isenberg (1986).

selectively seek and privilege information confirming their preexisting beliefs, while ignoring or minimizing contrary information.⁶³

Importantly, these two theories are not mutually exclusive, nor are they “mere theories.” Because both have significant empirical support⁶⁴, they are more accurately described as identifying *two primary causal mechanisms* of polarization.

Finally, a third theory, Self-Categorization Theory, also has some empirical support.⁶⁵ Self-Categorization Theory holds that when individuals are confronted with a risky or threatening outgroup, there is a pronounced human tendency to coalesce around the views of one’s self-identified in-group as a kind of protection-mechanism of solidarity against the threatening out-group.⁶⁶ This general mechanism is familiar from everyday life and history—as when Adolf Hitler used perceived threats to rally the German people behind his extreme Nazi ideology, as well as in polarized debates over moral issues today. For instance, anti-abortionists may be cast proponents of abortion as complicit in “genocide”⁶⁷, whereas proponents of abortion cast anti-abortionists as “waging a war on women.”⁶⁸ Finally, this mechanism appears to strengthen the more threatening the outgroup is perceived to be.⁶⁹

We can now provide several theoretical arguments that Discovery Model likely plays a significant role in group polarization. First, the Discovery Model appears to play directly into the phenomena described by Social Comparison Theory. Once again, Social Comparison

⁶³ Plous (1993): 233.

⁶⁴ For major confirming evidence of Social Comparison Theory, see Bray and Noble (1978), Myers and Bishop (1970), and Luhan et al (2009). For major confirming evidence of Informational Influence Theory, see Stoner (1961), Myers and Arenson (1972), Kaplan (1977), and Hinsz and Davis (1984).

⁶⁵ See Hogg et al (1990) and McGarty et al (1992).

⁶⁶ Hogg et al (1990).

⁶⁷ Cunningham (2009).

⁶⁸ Andrews et al. (2017).

⁶⁹ McGarty (1992).

Theory shows that people tend to seek approval of those they interact with, adopting progressively more extreme views to impress those in their group. As such, Social Comparison Theory predicts that if people cluster around opposing moral “discoveries”—if, for instance, some believe they have discovered abortion is wrong, whereas others believe they have discovered abortion is permissible—there will be a progressive tendency for each group’s members to adopt more extreme positions in order to impress members of their own group. Which, of course, is basically what we do see across a wide variety of moral issues. We see people cluster together in opposing moral groups—pro-abortion and anti-abortion groups, gun-control and gun-rights groups, pro-Israel groups and pro-Palestinian groups, etc.—with members on each side often “ramping up” their moral claims to impress fellow group-members.⁷⁰ Social Comparison Theory thus not only predicts that the Discovery Model likely plays a significant causal role in polarization. Social Comparison Theory and the Discovery Model together provide a potential explanation for why polarization is so pronounced on “judgmental” issues (specifically, moral issues).⁷¹ Because moral beliefs involve or are related to *reactive attitudes* such as anger and blame⁷², when members of opposing moral groups believe they have made opposing “moral discoveries” (viz. the Discovery Model), members of each group may adopt more extreme views in order to appeal to the reactive attitudes of members of their group (playing to their group’s anger, resentment, etc.).⁷³ Social Comparison Theory thus predicts not only that the Discovery Model likely plays a causal role in polarization, but that likely plays a prominent role.

⁷⁰ Tosi & Warmke (2016).

⁷¹ See e.g. Haidt (2012) and Isenberg (1986).

⁷² Strawson (1963).

⁷³ See Tosi & Warmke (2016): §2. For examples, see e.g. Cunningham (2009), Rostenberg (2014), and PoliticsUSA (2017).

Now turn to Informational Influence Theory, the theory which holds that polarization occurs by people in groups providing each other novel arguments and information that confirm their pre-existing beliefs, making them more amenable to even more extreme views. Here again, the Discovery Model appears to play directly into these phenomena. Informational Influence Theory predicts that that if people cluster around opposing moral “discoveries”—if, for instance, some believe they have discovered abortion is wrong, whereas others believe they have discovered abortion is permissible—people will tend to provide new arguments and information to members of their own favored moral group, making members of each group progressively more amenable to more extreme beliefs. Yet this too is broadly what we see in everyday life. In the abortion debate, for instance, we see many novel philosophical arguments generated on each side of the debate, with each side tending to emphasize the novel arguments for their own moral beliefs while ignoring or dismissively discounting arguments for the other side’s beliefs.⁷⁴ This is clearly not an isolated case, as people are known to cluster in political groups around different moral issues—groups which tend to expose their members to different information and arguments.⁷⁵ Further, as we saw earlier, Information Influence Theory predicts that polarization tends to be particularly pronounced for “intellectant” issues.⁷⁶ Because people who subscribe to the Discovery Model commonly treat moral matters as issues of intellectual debate—things to debate at dinner tables, on television, in university seminars, etc.—

⁷⁴ For instance, whereas anti-abortion websites such as Arthur (2001) and Pro-Life Perspective (2016) tend to emphasize philosophical arguments defending the moral status of human fetuses (viz. Marquis 2007 and Pojman 1998), pro-choice websites such as NLRC (2016) tend to emphasize arguments defending women’s rights to their bodies (viz. Thomson 1976).

⁷⁵ See e.g. Iyengar & Hahn (2009).

⁷⁶ Isenberg (1986).

Informational Influence Theory thus predicts that the Discovery Model likely plays a prominent role in polarization.

Finally, the Discovery Model also appears to play directly into the phenomena described by Self-Categorization Theory. Self-Categorization Theory predicts that group polarization tends to occur when an in-group is confronted by a threatening out-group. It is plain from everyday experience, however, that in-groups and “threatening outgroups” are often defined precisely in moral terms—in terms of “moral truths” people believe members of their group to have discovered. For instance, both sides of the abortion debate clearly find the other group threatening⁷⁷—and the same is clearly true across a variety of moral issues, including the Israeli-Palestinian conflict, gun-control, and so on. Insofar as human beings often do identify as members of moral groups—people who cluster around similar moral “discoveries”, viewing opposing groups as threatening out-groups—Social Categorization Theory also predicts that the Discovery Model plays a causal role in polarization.

3 An Anti-Polarizing Alternative? The Negotiation Model

In recent work, Muldoon and Arvan argue that many moral truths should be understood not as discoverable, but rather as the sort of thing that need to be *negotiated and renegotiated* on an ongoing basis.⁷⁸ Arvan argues that although a few regulative moral ideals (of coercion-minimization, mutual assistance, and fair bargaining) can be discovered through rational argumentation, all other moral truths—including how the above ideals should applied to applied ethical topics and politics—should be thought of not as discoverable through intuition, argument, or any other cognitive or affective process, but instead as *created by*

⁷⁷ Again, see Andrews et al. (2017) and Cunningham (2009).

⁷⁸ Arvan (2016) and Muldoon (2016).

interpersonal negotiation (and potentially renegotiated, if those affected remain unsatisfied with the norms that have been negotiated).⁷⁹ Muldoon's model, in contrast, focuses more squarely on political morality, arguing that because in a diverse society even basic moral norms are subject to fundamental disagreement, social and political principles in such a society should be settled and resettled on an ongoing basis according to an iterated bargaining process. Let us call this general approach the Negotiation Model of morality.

The basic idea behind the Negotiation Model is straightforward. Consider again the issue of abortion. The Discovery Model holds that there are preexisting facts about the morality of abortion (viz. its rightness, wrongness, etc.) that we can discover through intuition, argument, or some other cognitive or affective process. In contrast, the Negotiation Model holds that the moral status of abortion is literally indeterminate unless and until a social compromise has been arrived at, after which point abortion's moral status should be seen as *defined* by norms negotiated, in essence *settling* abortion's moral status via negotiated compromise (as in: "We have now negotiated a compromise that abortion is permissible in conditions A, B, and C, but impermissible in conditions X, Y, and Z"⁸⁰). On the Negotiation Model, as such, it is a meta- and normative-ethical mistake to form moral beliefs prior to social negotiation. People should instead withhold judgment on controversial moral issues, not forming moral beliefs on those issues until after clear public norms have been negotiated, after which point they should form beliefs in line with the negotiated norms (as in: "I now believe that abortion is morally permissible in conditions A, B, and C, but

⁷⁹ Arvan (2016).

⁸⁰ This, in brief, is why the Negotiation Model holds that moral *truth* is created via negotiated compromise: the outcome of compromise *comprises* facts about right, wrong, good, bad, etc. that did not exist previously, *establishing* socially created truth-conditions for moral claims.

impermissible in conditions X, Y, and Z, because these are the standards that have been publicly negotiated as a compromise”). Importantly, on the Negotiation Model, these publicly negotiated norms—and the moral beliefs they prescribe—are not mere “maxims” or rules to follow for some further moral aim (such as, say, utility-maximization). The norms instead express genuine *moral propositions* about the issue in question (viz. the moral permissibility or impermissibility of abortion) that individuals should *believe*.⁸¹

The Negotiation Model obviously raises many empirical and philosophical questions. First, is the model psychologically realistic? Can people really believe (for instance) that abortion’s moral status is indeterminate prior to public negotiation, and then come to believe that public negotiation settles its moral status? Second, is the model meta-ethically and normatively justifiable? Can negotiation truly *settle* the moral status of abortion? Because these are particularly broad issues that we cannot settle here, let us examine the Negotiation Model’s theoretical relationship to this paper’s topic—group polarization—as its implications in this narrow domain may be of some help in settling broader issues in future research.

Before proceeding further, it is worth recognizing that Negotiation Model is not a wholly new idea. Habermas famously defends a *discourse ethics*, according to which morality is (very roughly) a matter of seeking mutual understanding.⁸² Similarly, Rorty argued that moral and social institutions are better thought of as, “experiments in cooperation rather than as attempts to embody a universal and ahistorical order.”⁸³ A bit less strongly, Ackerman argues that a morally reflective person cannot permissibly cut themselves off

⁸¹ I thank an anonymous reviewer for suggesting this clarification.

⁸² Habermas (1981, 1995).

⁸³ Rorty (1991): 196.

from political dialogue.⁸⁴ Finally, although Rawls argues that principles of justice can be discovered through reflecting on a hypothetical thought-experiment (the ‘original position’), there are also undercurrents of discourse ethics (and the Negotiation Model) in his work. To see how, consider that Rawls maintains in *A Theory of Justice* that, ‘justification is argument addressed to those who disagree with us...Being designed to reconcile by reason, justification proceeds from what all *parties to the discussion* hold in common.’⁸⁵ Rawls’ recognition that justification involves discussion with others who at least initially disagree—and that discussion should seek a kind of common ground—appears to involve some kind of commitment to *negotiating* requirements of justice. And indeed, Rawls adds in his later work *Political Liberalism* that, ‘While I view it [the two principles of justice as fairness] as the most reasonable...I would simply be unreasonable if I denied that there were other reasonable conceptions...Any conception [of justice] that meets the criterion of reciprocity and recognizes the burdens of judgments...is a candidate [for a reasonable conception of justice].’⁸⁶ This passage, particularly Rawls’ recognition that there could be *other* theories of justice that reasonable people might converge on—suggests that, for Rawls, although some regulative ideals of political morality (reciprocity and burdens of judgment) can be recognized through reason, the question of which *particular* theory is most reasonable is one we must negotiate with others through mutual dialogue. If this is the right way to understand Rawls, then although his theory may appear to superficially conform to the Discovery Model (viz. the idea that we can discover his two principles of justice as ones that would govern a fully just society), at a more subtle level Rawls may well be a proponent of a form of the

⁸⁴ Ackerman (1989).

⁸⁵ Rawls (1999): 508.

⁸⁶ Rawls (1993): xlix.

Negotiation Model. Consequently, Arvan's and Muldoon's theories can perhaps be better thought of as providing new defenses and developments of an approach to moral and political philosophy that has already had substantial influence (though, as we have seen, the Discovery Model has been far more dominant, both in everyday life among laypeople and academic moral philosophers). If this article is correct, group polarization may provide new reasons for philosophers and laypeople to take this alternative tradition more seriously.

Because as we have seen there is more than one theory of *how* morality may be fundamentally a matter of negotiation, let us define 'the Negotiation Model' of morality very broadly, as follows:

The negotiation model of meta-ethics: aside (perhaps) from some *general regulative moral ideals* which may be discovered by rational argumentation (such as ideals of coercion-minimization and mutual assistance), moral truths do not exist to be discovered (or justified moral beliefs to be formed) by intuition, argument, or any other unilateral cognitive or affective process, but are instead *created* by multi-lateral, interpersonal processes of moral negotiation involving all those affected *seeking to arrive at a compromise agreement* guided by the aforementioned regulative ideals.

The negotiation model of moral-belief formation: people who endorse the negotiation model of meta-ethics will tend to *seek common ground* with others who share relevant regulative moral ideals; believe that morality *requires* moral negotiation as such; and *form moral beliefs* only after negotiating compromises, in line with whatever norms have resulted from compromise agreement.

Because the model is complex, allow me to spell out each component.⁸⁷ “Moral negotiation”, as I wish to understand it, is a term of art intended to cover any and all forms of interpersonal human behavior (such as conversation, voting, and so on) that involve those with interests in a given moral issue (abortion, etc.) obeying certain discursive rules—rules that at the very least include a commitment to (A) *certain regulative moral ideals* as background beliefs and motivations, (B) *suspending moral judgment* on the issue in question before interpersonal agreement is reached, and (C) *seeking a compromise agreement* on moral norms for the issue in question (e.g. abortion) with others who plausibly satisfy conditions (A) and (B). Allow me to illustrate.

As mentioned earlier, on my account moral truths on controversial moral issues (e.g. abortion) should be understood in terms of the *outcome* of a compromise agreement between all those who plausibly share certain regulative ideals (of coercion-minimization, mutual assistance, etc.), but who may have different preferences and priorities on the issue in question. The example of abortion is instructive here. On my account, a major reason why the morality of abortion remains such a divisive issue is because individuals on both sides of the debate plausibly share relevant regulative moral ideals. Anti-abortion advocates, for example, typically claim that fetuses have a “right to life”—a claim clearly intended to *help fetuses*, protecting them from having their lives *coercively* ended. Pro-abortion advocates, on the other hand, claim women have a “right to choose”—a claim clearly intended to *help women*, protecting their reproductive choices from being *coercively reduced*. Both sides are thus plausibly motivated by regulative ideals of coercion-minimization and assisting others. The primary difference between the two sides—on my version of the Negotiation Model—

⁸⁷ I thank an anonymous reviewer for inviting me to explicate these components.

occurs at the level of moral priorities: anti-abortion advocates currently think the rights of the fetus should *outweigh* the rights of women, whereas pro-abortion advocates think the rights of women should take priority. On my version of the Negotiation Model, because both sides share relevant regulative ideals but have different priorities, they have a duty to negotiate a compromise: it is *wrong* for both sides to think that they can “discover” moral correctness of their own preferred view (e.g. “Abortion is wrong!”) via intuition, argument, or any other process of discovery.

On the Negotiation Model, then, telling someone who cares about protecting fetuses or women they have the “wrong” view on abortion is fundamentally antithetical to morality—as this violates the Negotiation Model’s discursive requirement to suspend judgment prior to the outcome of negotiation. In contrast, the following conversational moves would be legitimate: “You and I both want to protect people from coercion. You want to protect fetuses from death, and I want to protect women’s ability to make family planning choices. You want women to either remain abstinent, or else accept the consequences of their sexual choices—including unintended pregnancy. However, I find this problematically coercive, as it is very difficult for people to remain abstinent, and so on. Because you and I disagree over these matters, let us seek a middle ground—one that may not fully satisfy either of us, but will at least enable you and I to arrive at a compromise, one involving substantial protection of fetal life for you, protection of women’s choice for me.”

What kind of negotiated compromise might emerge from such a process? One obvious possibility—but not the only possible one—is this: because prevailing scientific knowledge indicates fetuses first become sentient between 18 to 25 weeks of gestation⁸⁸, both sides of

⁸⁸ See Tawia (1992) and Koch (2009).

the abortion debate could (if they were willing to conform to the Negotiation Model) arrive at a compromise agreement that (i) early abortion prior to fetal sentience is morally permissible, (ii) abortion after fetal sentience is normally impermissible, except perhaps in cases of rape, incest, or danger to the mother's life, and finally (iii) members of society share a duty to devote ample social resources to provide sexually-active women with ready access to family-planning resources (including access to affordable early-term abortion) to prevent abortion after fetal sentience.

Such a compromise would almost certainly not fully satisfy many parties to the abortion debate—in part, I think, because the Discovery Model is so deeply entrenched in how people think about moral issues (viz. “But abortion is *murder!*”), but also because of the very nature of compromise (which requires “give and take”). Still, as uncomfortable as compromise may be, we should not summarily dismiss such accounts. Importantly, because people might not be wholly satisfied with a given compromise, compromises following the Negotiation Model would plausibly leave many matters open to *renegotiation*. If, for instance, the above compromise on abortion had serious negative effects on women's lives (e.g. by requiring single women to bear children if the father dies late in pregnancy), individuals on the “pro-choice” side of the debate could bring that new information to bear publicly in the aim of renegotiating abortion norms (as in: “I know we have currently agreed that abortion is permissible in cases X, Y, and Z. However, because this standard of permissibility is having negative effects on women, I would like us to consider a new compromise on different standards of permissibility”). The Negotiation Model thus entails—plausibly, I believe—that moral truths on controversial issues can literally *evolve* as people lobby for and negotiate new compromises.

It might seem hard to imagine many people ever accepting the Negotiation Model, as it would require us to give up many moral convictions that we may believe very deeply. Indeed, the Negotiation Model's psychological model of moral-belief formation might turn out to be difficult or even impossible for people to reliably conform to, if Haidt's Social Intuition Model (SIM) of moral-belief formation is correct.⁸⁹ Further, some may worry that the Negotiation Model would have undesirable moral consequences, such as requiring gays and lesbians to negotiate on same-sex marriage—compromises that might set back the gay liberation movement.⁹⁰ These are important questions worth investigating in more detail in future research. Indeed, it is worth noting that the Negotiation Model's relations to group polarization could potentially help address some of the concerns just raised. Suppose, in line with my arguments to come in this paper, empirical investigation found the Negotiation Model to prevent polarization and foster cooperative resolutions to moral controversies. Suppose, further, that empirical studies found cooperative resolutions to moral debates to better advance the cause of liberation groups, on average, relative to the Discovery Model—which the Negotiation Model might do in at least two ways: (1) by *delegitimizing* certain traditional forms of moral argument (e.g. appeal to religious scripture or traditional practices to make 'moral discoveries' about same-sex marriage), and (2) *undermining* the formation and effectiveness of reactionary social movements (e.g. anti-gay groups) whose members oppose liberation movements. As such, future research might either dispel or

⁸⁹ See Haidt (2012). However, it is important to note here that empirical data supporting the SIM model have been collected under prevailing social conditions—in which most people appear to tacitly or explicitly accept the Discovery Model. Consequently, it is an open question whether a “paradigm shift” in the direction of the Negotiation Model might substantially change how individuals form moral beliefs.

⁹⁰ I thank an anonymous reviewer for raising these concerns.

confirm the above worries about the Negotiation Model. However, because we cannot settle these issues here, let us instead examine the model's relationship to polarization.

Let us begin with Social Comparison Theory: the theory which holds that group polarization results from people adopting more extreme views to *impress* fellow group members. The Negotiation Model promises to undercut this polarizing mechanism in at least two ways: by (A) discouraging people from adopting first-order moral beliefs (e.g. "Abortion is wrong!") prior to negotiating, and (B) encouraging people to see answers to controversial moral questions as *created* through negotiated compromise. These two discursive elements of the Negotiation Model promise to defuse polarizing mechanisms of social comparison at "step 1": it would lead people *have no settled moral beliefs* for groups to cluster around (e.g. "Abortion is wrong") in the first place. On the contrary, it would plausibly give people an *anti-polarizing* ideal to cluster around despite their differences: the ideal of negotiating compromises with people with different priorities who share relevant moral ideals. Importantly, existing research already indicates that group organization around such a cooperative ideal does indeed mitigate polarization and promote cooperation.⁹¹ As such, the Negotiation Model theoretically promises to harness the forces that Social Comparison Theory identifies as responsible for group polarization to produce the very opposite: progressive convergence on a cooperative norm of negotiated compromise.

Now consider Informational Influence Theory, the theory which holds that group polarization is generated precisely by people seeking out and attending selectively to information that confirms their preexisting belief. The Negotiation Model promises to undermine the social psychological forces responsible for group polarization here as well.

⁹¹ Brewer (1996).

When people have preexisting first-order moral beliefs on a given side of an issue, as on the Discovery Model—such as the preexisting belief that abortion is wrong, or alternatively, that women have a right to abortion—the phenomenon of informational influence leads *each side* to attend to information in support of their preexisting moral beliefs, making each side more amenable to more extreme beliefs. The Negotiation Model, on the other hand, holds that people should *not have* opposing first-order moral beliefs in the first place, but should instead conceive answers to moral questions (such as the morality of abortion) as *created* through negotiated compromise. But now if people increasingly held this kind of first-order moral belief—the belief that answers to controversial moral questions must be negotiated—then Informational Influence Theory predicts that people would become progressively more inclined to seek out and attend selectively to their belief that *moral answers must be created cooperatively* through negotiation. Consequently, the Negotiation Model theoretically promises to harness the phenomena described by Informational Influence Theory to *prevent* polarization and promote cooperation.

Finally, consider Social-Categorization Theory, the theory which holds that group polarization results from the development of *in-groups* which then treat out-groups as threatening. As we have seen, in-groups and out-groups often form around divisive moral “discoveries.” For example, whereas anti-abortionists often cast defenders of abortion as a threatening out-group (“They are baby-killers!”), defenders of abortion often cast anti-abortionists the very same way (“They want to take away women’s rights!”). The Negotiation Model once again promises to halt this polarizing force at “step 1.” Insofar as it (A) holds that people should not have settled moral beliefs on issues prior negotiation, and (B) should be willing to negotiate compromises with those with different priorities, the Negotiation Model

would theoretically *prevent the formation* of divisive in-groups and out-groups, instead promoting the development of a *cooperative in-group*: people who have different priorities on controversial moral issues (abortion, gun control, etc.), but who are nevertheless unified around anti-polarizing ideals of negotiation and compromise (something which, again, has indeed been found to promote cooperation⁹²). Finally, Social-Categorization Theory does plausibly predict that the Negotiation Model would generate certain types of polarization—namely, polarization between those who accept the Negotiation Model and those who accept the Discovery Model (who might indeed regard each other as threatening “enemies”), as well as polarization with those who reject relevant regulative ideals (e.g. racists, sexists, etc.). However, while Social-Categorization Theory plausibly predicts that there would be polarization between these groups—with each potentially treating the others as threatening out-groups—these would not obviously be bad forms of polarization according to the Negotiation Model, as the model itself suggests that we should *not* tolerate the Discovery Model or false regulative ideals.

In sum, all three theories of polarization predict that the Negotiation Model is likely to substantially reduce polarization relative to the Discovery Model, and perhaps even generate its opposite: a progressive willingness to cooperate and compromise.

4 Normatively Evaluating Polarization and Models of Morality

This article cannot definitively settle how to normatively evaluate polarization or the two models of morality under discussion. That would require answers to broad meta- and normative- ethical questions well beyond the scope of a single article. We can, however, outline some plausible methods for investigating these matters.

⁹² Brewer (1996).

First, it is worth recognizing that traditional models of morality—the kinds identified earlier as falling into the Discovery Model—might be used to justify *adopting* the Negotiation Model of meta-ethics and moral-belief formation on *meta-ethically self-effacing first-order moral grounds*, due to the models' hypothesized effects on polarization. Allow me to explain.

Consider standard act-utilitarianism. As mentioned earlier, act-utilitarianism has been traditionally understood along the lines of the Discovery Model. Utilitarian philosophers aim to *discover* what maximizes utility (in trolley cases, torture, freedom of expression, and so on), arguing that we should *adopt moral beliefs* on the basis of those discoveries (e.g. believe torture is wrong if torture is discovered not to maximize utility). However, utilitarianism could in principle entail—if the Discovery Model is found to produce polarization, and polarization is in turn found not to maximize utility—that we should *reject* the Discovery Model in favor of the Negotiation Model *as a means to maximizing utility*. After all, if utilitarianism is true, then we should *morally* evaluate polarization and meta-ethical models of moral truth, justified belief, and moral-belief formation wholly by reference to *social utility*. To see how, suppose this paper's primary theoretical predictions are verified by future research—that the Discovery Model turns out to substantially cause, and Negotiation Model substantially reduce, group polarization. Suppose that further research then found due to their respective effects on polarization, our subscribing to the Negotiation Model (of moral truth, justified belief, and moral-belief formation) would *cause greater social utility* than the Discovery Model (perhaps by the former fostering better social cooperation and less divisiveness). If this is what future empirical science found, we would have utilitarian grounds for *subscribing* to the Negotiation Model over the Discovery Model (even if, ultimately, the Discovery Model is *true* meta-ethically).

Are there any reasons to believe that utilitarianism might actually have these implications? There are *prima facie* reasons to think it might. First, commitment to ideals of negotiation has already been found to foster productive cooperation.⁹³ Second, while the Negotiation Model allows for some forms of polarization—forms caused by fighting racism, sexism, and other values inconsistent with the model’s regulative ideals—these specific forms of polarization are *prima facie* utility-conducive. Because racists, sexists, and so on, aim to deny happiness to entire classes of people, fighting against such values (which the Negotiation Model requires) is plausibly conducive to long-term social utility, even if it involves polarization. Indeed, this seems evident from civil rights movements, which are often polarizing yet plausibly lead to greater long-run utility by overthrowing oppression. Finally, there are *prima facie* reasons to believe that other forms polarization—including kinds plausibly caused by the Discovery Model—are utility-*diminishing*. For, as we have seen, polarization in general is known to have many *prima facie* negative social effects, including increasing social mistrust, selfishness, extremism, and violence. We plausibly witness these negative effects on a continual basis, across a wide variety of domestic and international contexts, such as the Israeli-Palestinian conflict, where polarization has been shown to undermine the ability of the opposing sides to forge a stable, peaceful resolution.⁹⁴ Another obvious example is domestic political affairs in the United States today, where increasing polarization⁹⁵ has been accompanied increasing public dissatisfaction with Congress⁹⁶ and the overall direction of the country.⁹⁷ There are of course many complex empirical questions

⁹³ Brewer (1996).

⁹⁴ Kelman (1982).

⁹⁵ Dimock *et al* (2014) and Yardi & Boyd (2010).

⁹⁶ Gallup (2017).

⁹⁷ RealClearPolitics (2017).

here. The relevant point is merely that there are *prima facie* reasons to believe that the Negotiation Model may be more utility-conducive than the Discovery Model due to each model's hypothesized effects on polarization.

Second, while as we saw earlier Kantianism has been traditionally interpreted in line with the Discovery Model—as Kantian philosophers generally suppose we can *discover* which maxims are universalizable or respect humanity as an end-in-itself through philosophical argumentation—it is *possible* that Kantianism could morally require us to *reject* the Discovery Model in favor of the Negotiation Model on first-order moral grounds. Consider for example Christine Korsgaard's influential "practical contradiction interpretation" of Kant's categorical imperative, which holds that a maxim of actions fails to satisfy Kant's principle—and is therefore morally wrong to act upon—if the universal adoption of the maxim as a "standard procedure" would *undermine* the successful pursuit of the maxim's intended goal.⁹⁸ As illustrated above with utilitarianism, the kinds of polarization-producing actions the Negotiation Model plausibly allows—polarization in fighting racism, sexism, and so on—may well be *necessary* for the effective pursuit of their end-goal: overthrowing oppression. Because, as many civil rights movements suggest, overthrowing oppressive values may require polarizing activities (protests, civil disobedience, etc.), forms of polarization caused by the Negotiation Model's requirement to not negotiate with inherently-oppressive values may well pass Korsgaard's practical contradiction test. Finally, there are *prima facie* reasons to believe that other forms of polarization—including kinds potentially caused by the Discovery Model—generally make it more difficult for agents to effectively pursue their goals, thus plausibly failing Korsgaard's

⁹⁸ Korsgaard (1996): 92-4.

test. For example, polarization in the Israeli-Palestinian conflict appears to have undermined the ability of both sides to effectively pursue their goals.⁹⁹ Similarly, political polarization in the United States appears to have led to consistent frustrations on both sides of the political aisle, as evidenced again by low congressional approval ratings¹⁰⁰ and dissatisfaction with the direction the country is heading.¹⁰¹ Thus, while there are once again many complex questions—regarding which kinds of polarization undermine the successful pursuit of goals—there are *prima facie* reasons to believe that adopting the Negotiation Model as a maxim (for understanding moral truth, justified moral-beliefs, and forming moral-beliefs) might *satisfy* the categorical imperative whereas adopting the Discovery Model might fail.

Third, virtue-ethical frameworks—which have also been traditionally interpreted in line with the Discovery Model (viz. our task is to discover how to live virtuously)—could in an analogous fashion *require* us, as a first-order moral requirement, to reject the Discovery Model in favor of the Negotiation Model. For consider Rosalind Hursthouse’s definition of morally right action in terms of what a *virtuous agent* would do in the circumstances.¹⁰² On the one hand, forms of polarization allowed by the Negotiation Model (e.g. polarization caused by protesting racism or sexism) plausibly express *moral virtue*. If polarization is necessary for overthrowing oppression—as the history of civil rights movements plausibly suggest—and the virtuous person just is a person *committed* to overthrowing oppression (which seems plausible as well), then there are *prima facie* reasons to believe that the virtuous person would engage in forms of polarization allowed by the Negotiation Model. On

⁹⁹ Kelman (1982).

¹⁰⁰ Gallup (2017).

¹⁰¹ RealClearPolitics (2017).

¹⁰² Hursthouse (1999).

the other hand, polarization more generally—including kinds of moral polarization theoretically caused by the Discovery Model (viz. the Israeli-Palestinian conflict, contemporary politics, etc.)—has been linked to things commonly deemed *moral vices*: selfishness, mistrust, prejudice, extreme political decisions, group violence, and so on. Consequently, although once again many complex questions arise—about which forms of polarization are virtuous or vicious—there are *prima facie* reasons to believe that virtue ethics may also support subscribing to the Negotiation Model over the Discovery Model.

Now, to be sure, there is something philosophically awkward—and, I think, ultimately problematic—about defending the Negotiation Model in these ways. For notice: on all of the above frameworks, morality is still a matter of *discovering* (viz. utility, universalizability, virtue) whether we ought to adopt the Negotiation Model on “moral grounds.” As such, all such views are in a deep way *metaethically self-effacing*, holding (*qua* Discovery Model) that moral truths really are ultimately *discoverable*, but that to *fulfill* our moral duties we need to act (psychologically) *as though* the Negotiation Model were true. While such a self-effacing view—that first-order morality requires us to *believe false things* about metaethics (viz. the Negotiation Model)—may ultimately be correct, it is for all that an awkward view, as all self-effacing philosophical views are.

Fortunately, there are *non-self-effacing* ways of defending the Negotiation Model as a *correct* meta-ethical view partly on the basis of the model’s relations to polarization. For example, Muldoon argues that seeing morality in terms of bargaining is advantageous for social cooperation, as parties willing to bargain may be more willing to cooperate with others who disagree than those who ‘stand on principle.’ Notice that Muldoon’s view sits nicely with the present paper’s hypotheses regarding polarization: namely, that at least some forms

polarization may be *normatively bad* because of their tendency to make social cooperation more difficult (though perhaps other forms of polarization—those overthrowing oppressive systems such as slavery—may *improve* social cooperation). Somewhat similarly, Arvan argues it is rational to conceptualize morality as requiring negotiation on grounds of *diachronic rationality*. In brief, Arvan argues that because individuals make choices that affect themselves in the future, it is rational to choose in ways that advance the interests of one's present *and* future selves.¹⁰³ But, Arvan argues, because the future (particularly the distant future) is deeply uncertain, the most rational way to choose in the present may be to act in ways that *as many of one's possible future selves* can endorse as possible. Consequently, Arvan argues, because some of one's possible future selves may be purely self-interested, whereas other possible future selves care about the interests of others for a variety of reasons—including self-interest, but also transformative empathy—justifying one's actions to one's present and possible future selves requires *negotiating a middle ground* between one's possible egoistic interests and the interests of other beings (things Arvan suggests requires acting on regulatory goals of coercion-minimization, mutual assistance, and fair bargaining in negotiation). While on Arvan's model some forms of polarization may be normatively rational (namely, polarization that aims to overthrow ideals inconsistent with regulatory ideals of coercion-minimization and mutual assistance), other forms of polarization (forms that undermine compromise between individuals who *share* relevant regulatory ideals) are normatively irrational.

In short, Muldoon and Arvan both contend morality should be fundamentally reconceived in terms of negotiation on grounds of *rationality* (*viz.* cooperating despite

¹⁰³ Arvan (2016): ch. 2.

disagreement for Muldoon, diachronic individual rationality more generally for Arvan). Because some (if not all) forms of group polarization plausibly undermine social cooperation or the goals of one's possible future selves, Muldoon and Arvan both offer additional theoretical resources for normatively evaluating group polarization and, by extension, comparatively evaluating the Discovery and Negotiation models of morality.

5 Avenues for Future Research

This paper's argument raises many empirical questions. First, do people generally conform to the Discovery Model, or do some of us already at least tacitly endorse some form of the Negotiation Model? Although people often seem to form moral beliefs as "unilateral" discoveries, people also sometimes seem willing to negotiate moral compromises. Second, can people realistically form beliefs in the manner prescribed in the Negotiation Model, suspending judgment on moral issues until after compromises are publicly negotiated? Third, does the Discovery Model actually cause group polarization in the ways identified in §2, and would adopting the Negotiation Model prevent polarization and promote productive compromise in the ways hypothesized in §3? Finally, there is the interesting question—which due to space-constraints are beyond this paper's scope—of how *non-cognitivist* moral views (which deny the existence of moral truth altogether) relate to polarization relative to the two models discussed here.¹⁰⁴

These are all important questions that could be examined utilizing common empirical methodologies. First, cross-sectional survey studies might measure whether and to what extent people accept the Discovery Model (viz. survey items like, "I adopt my moral beliefs on the basis of intuition", "I have settled beliefs about the morality of abortion", etc.); whether

¹⁰⁴ I thank an anonymous reviewer for encouraging me to note this.

and to what extent people accept the Negotiation Model (viz. items like, “I am willing to compromise with others on moral issues”); and finally, whether and to what extent respondents engage in polarizing behaviors. Cross-sectional studies of this sort might enable researchers to determine whether and to what extent the Discovery Model and Negotiation Model are related to polarization-causing behaviors (of the sort identified by the three theories of polarization). Second, longitudinal studies might be used to establish causation. For example, a longitudinal study might have a control group who engage in ‘normal’ moral behavior (viz. the Discovery Model) alongside experimental groups whose participants might be prompted, several times per day when engaging in social media use, to “withhold judgment” or otherwise conform to the Negotiation Model’s discursive rules. Longitudinal studies of this sort might enable researchers to determine whether prompting people to conform to the Negotiation Model lessens polarizing behaviors relative to the Discovery Model. Finally, experimental interventions might be studied under laboratory conditions, with small groups of test-subjects *debating* a moral issue, where in the control condition respondents act “normally” (viz. the Discovery Model) and in the experimental condition respondents are prompted to conform to the Negotiation Model.

Finally, this paper raises many philosophical questions. First, there is the question of which of the two models of morality—the Discovery Model or Negotiation Model—is most defensible *meta-ethically* (viz. moral semantics, etc.). Second, this paper’s theoretical arguments raise important *normative ethical* questions about the desirability of polarization and different models of morality and moral-belief formation. For example, because homogenous groups may outcompete heterogeneous groups¹⁰⁵, some forms of group

¹⁰⁵ See e.g. Durkheim [1893] & Haidt (2012).

polarization may be socially adaptive. Similarly, although exposing people to a variety of opinions (as the Negotiation Model suggests) appears to promote moderation in group opinions and behaviors¹⁰⁶, there is also some evidence that it can promote apathy.¹⁰⁷ Because the social effects of the Discovery Model and Negotiation Model are likely to be complex, understanding them better is likely to be of substantial philosophical importance.

Conclusion

Laypeople and philosophers tend to treat moral truths or justified moral beliefs as “discoverable” through intuition, argument, or other cognitive or affective process. However, we have seen that there are strong theoretical reasons—based on three empirically-supported theories of group polarization—to believe this Discovery Model of morality is a likely cause of polarization: a social-psychological phenomenon known to have a wide variety of disturbing social effects. We then saw that there are complementary theoretical reasons to believe that a Negotiation Model of morality might not only mitigate polarization but actually foster its opposite: an increasing willingness for to work together to arrive at compromises on moral controversies. While this paper does not prove the existence of the hypothesized relationships between the Discovery Model, Negotiation Model, and polarization, it demonstrates that there are ample theoretical reasons to believe that such relationships are likely and worthy of further empirical and philosophical research.

¹⁰⁶ Johnston & Macrae, 1994; Stroud, 2010

¹⁰⁷ Huckfeldt, Morehouse Mendez, & Osborn, 2004

References

Abrams, D.; M. Wetherell; S. Cochrane; M.A. Hogg; J.C. Turner (1990). 'Knowing what to think by knowing who you are: Self-categorization and the nature of norm formation, conformity and group polarization', *British Journal of Social Psychology* 29 (2): 97-119.

Ackerman, B. (1989). 'Why Dialogue?', *Journal of Philosophy* LXXXVI(1): 5-22.

Adams, S. (2017). 'Country Musician Caleb Keeter Changed His Mind on Gun Control After Getting Caught in the Las Vegas Shooting', *Slate*, Oct 3 2017: <https://www.google.com/amp/amp.slate.com/blogs/browbeat/2017/10/03/country-musician-changes-mind-on-gun-control-after-las-vegas-shooting.html>, accessed on Oct 6, 2017.

Andrews, B., Levintova, H., Liss-Schultz, N. & P.R. Lockhart (2017). 'Trump Brought the War on Women Mainstream in His First 100 Days', *Mother Jones*, <http://www.motherjones.com/politics/2017/04/first-100-days-reproductive-rights-abortion-trump>, accessed on 6/5/2017.

Arthur, J. (2001). 'Personhood: Is a Fetus a Human Being', *Pro Choice Action Network*, <http://www.prochoiceactionnetwork-canada.org/articles/fetusperson.shtml>, accessed 3/20/16.

Aristotle, *Nicomachean Ethics*, T. Irwin (trans.) (1999). *Aristotle: Nicomachean Ethics*, 2nd Ed. (Indianapolis: Hackett).

Arvan, M. (2016). *Rightness as Fairness: A Moral and Political Theory* (London: Palgrave MacMillan).

Audi, R. (2015). 'Intuition and Its Place in Ethics', *Journal of the American Philosophical Association* 1(1): 57-77.

Baldassarri, D., & Gelman, A. (2008). 'Partisans without constraint: Political polarization and trends in American public opinion', *AJS: American journal of sociology*, 114(2), 408.

Barcalow, E. (2007). *Moral Philosophy: Theories and Issues*, 4th ed. (Belmont, CA: Thomson Wadsworth).

BBC (2009a). 'Sanctity of Life: Islamic Teachings on Abortion'.
http://www.bbc.co.uk/religion/religions/islam/islamethics/abortion_1.shtml,
accessed 12/28/2018.

----- (2009b). 'Hinduism and Abortion'.
http://www.bbc.co.uk/religion/religions/hinduism/hinduethics/abortion_1.shtml,
accessed 12/28/2018.

BBC Bitesize (n.d.). 'What Does Judaism Teach About Abortion', <https://www.bbc.com/bitesize/guides/zt6ncdm/revision/4>, accessed 12/28/2018.

Black, P.; B. Wedeman, & J. Mullen (2015). 'Israeli Palestinian Conflict: Two Viewpoints; One Tragic Outcome', *CNN*, <http://www.cnn.com/2015/10/21/middleeast/israeli-palestinian-root-of-violence>, accessed 6/20/16.

Bray, R. M. & A. M. Noble (1978). 'Authoritarianism and decisions of mock juries: Evidence of jury bias and group polarization', *Journal of Personality and Social Psychology* 36 (12): 1424-1430

Brewer, M. B. (1996). 'When contact is not enough: Social identity and intergroup cooperation', *International journal of Intercultural relations*, 20(3-4), 291-303.

Brezenoff, D. (2018). 'Ten Arguments Against Gun Control—And Facts That Prove Them Wrong.' *Medium*, <https://medium.com/@dbrezenoff/ten-arguments-against-gun-control-and-facts-that-prove-them-wrong-ce20eb33633d>, accessed 12/28/2018.

Brink, D. (1989). *Moral Realism and the Foundations of Ethics* (Cambridge: Cambridge University Press).

Cunningham, G. (2009). 'Is Abortion Genocide?', *Canadian Center for Bio-Ethical Reform*, <https://www.endthekilling.ca/training/classroom/genocide>, accessed 6/2/2017.

Dancy, J. (2013). 'Moral Particularism', in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), <http://plato.stanford.edu/archives/fall2013/entries/moral-particularism/>.

----- (1986). 'Two Conceptions of Moral Realism', *Proceedings of the Aristotelian Society, Supplementary Volume 60*: 167-87.

Daniels, N. (2013). 'Reflective Equilibrium', in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2013 Edition), <http://plato.stanford.edu/archives/win2013/entries/reflective-equilibrium/>, accessed 6/20/16.

Dimock, M., Doherety, J.K., & Oates, R. (2014). 'Political Polarization in the American Public: How increasing ideological uniformity and partisan antipathy affect politics, compromise, and everyday life.' *Pew Research Center*.

DuCharme, W. M. (1970). 'Response bias explanation of conservative human inference.' *Journal of Experimental Psychology*, 85(1), 66.

Durkheim, E. [1893]. *The Division of Labor in Society*, G. Simpson (trans.) (New York: the Free Press).

Einstein, A. (1905) "On the electrodynamics of moving bodies", *Annalen Der Physik*: 891-921.

Eternal Word (1996). 'What Does the Bible Say About Abortion? What About the Early Christians?', *Eternal Word Television Network*, available at: <http://www.ewtn.com/library/prolife/bibsayab.txt>, accessed 12/28/2018.

Feminist eZine (n.d.). 'Abortion: It's Every Woman's Right to Choose.' <http://www.feministezine.com/feminist/anarchist/Anarchist-Perspective-Abortion.html>, accessed 12/28/2018.

Finlay, S. (2007). 'Four Faces of Moral Realism', *Philosophy Compass* 2: 820-49.

Gallup (2017). 'Congress and the Public', www.gallup.com/poll/1600/congress-public.aspx, accessed August 24, 2017.

----- (2013). 'Older Americans' Moral Attitudes Changing', <http://news.gallup.com/poll/162881/older-americans-moral-attitudes-changing.aspx>.

Habermas, J. [1992]. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy* (Wiley & Sons, 2015).

----- (1981). *The Theory of Communicative Action* (Beacon Press).

Haidt, J. (2012). *The righteous mind: why good people are divided by politics and religion* (New York: Vintage Books).

---- (2001). 'The emotional dog and its rational tail: a social intuitionist approach to moral judgment', *Psychological review*, 108(4), 814.

Higgs, P.Q. (1964). 'Broken Symmetries and the Masses of Gauge Bosons'. *Physical Review Letters* 13 (16): 508-509

Hinsz, V.B.; J.H. Davis (1984). 'Persuasive Arguments Theory, Group Polarization, and Choice Shifts', *Personality and Social Psychology Bulletin* 10: 260-268.

Hoffrage, U. (2004). 'Overconfidence', in Pohl, R. (Ed.). (2004). *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory* (Psychology Press).

Hogg, M.A.; Turner, J.C.; Davidson, B. (1990). 'Polarized norms and social frames of reference: A test of the self-categorization theory of group polarization', *Basic and Applied Social Psychology* 11: 77-100.

Hook, J. (2016). 'Why I Changed My Mind About Gay Marriage', <http://www.joshuahook.com/changed-mind-gay-marriage>, accessed on Oct 6, 2017.

Huckfeldt, R., Mendez, J.M., & Osborn, T. (2004). 'Disagreement, ambivalence, and engagement: The political consequences of heterogeneous networks', *Political Psychology*, 25(1), 65-95.

Hursthouse, R. (1999). *On Virtue Ethics* (Oxford: Oxford University Press).

Isenberg, D.J. (1986). 'Group Polarization: A Critical Review and Meta-Analysis', *Journal of Personality and Social Psychology* 50 (6): 1141-1151.

Iyengar, S. & Hahn, K.S. (2009). 'Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use', *Journal of Communication* 59(1): 19-39.

Johnston, L.C. & C.N. Macrae. 'Changing social stereotypes: The case of the information seeker', *European Journal of Social Psychology* 24.5 (1994): 581-592.

Joyce, R. (2015). 'Moral Anti-Realism', in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), <http://plato.stanford.edu/archives/sum2015/entries/moral-anti-realism/>.

----- (2007). *The Myth of Morality* (Cambridge, UK: Cambridge University Press).

Kant, I. (1797). *The Metaphysics of Morals*, in M.J. Gregor (ed.), *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy* (Cambridge: Cambridge University Press, 1996): 363-603.

---(1785). *Groundwork of the Metaphysics of Morals*, in M.J. Gregor (ed.), *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy* (Cambridge: Cambridge University Press, 1996): 38-108.

Kaplan, M.F. (1977). 'Discussion Polarization Effects in a Modified Jury Decision Paradigm: Informational Influences', *Sociometry* 3 40(3): 262-271.

Katsafanas, P. (2011). 'Deriving Ethics from Action: A Nietzschean Version of Constitutivism', *Philosophy and Phenomenological Research* 83(3): 620-60.

Kaufman, D. (2017). 'August 15, 2017 at 12:31pm comment at *Daily Nous*', <http://dailynous.com/2017/08/15/white-supremacists-charlottesville-philosophy-classroom>, accessed August 24, 2017.

Kelman, H.C. (1982). 'Creating the conditions for Israeli-Palestinian negotiations.' *Journal of Conflict Resolution* 26(1): 39-75.

Keseman, B. (n.d.). 'Sanctity of Human Life Sunday: A Sermon Suggestion'. <http://www.lcms.org/Document.fdoc?src=lcm&id=206>, accessed 12/28/2018.

Koch, C. (2009). 'When Does Consciousness Arise in Human Babies?', *Scientific American*,
<https://www.scientificamerican.com/article/when-does-consciousness-arise/>,
accessed 6/1/2017.

Korsgaard, C.M. (2009). *Self-Constitution: Agency, Identity, and Integrity* (Oxford: Oxford University Press).

----- (2008). *The Constitution of Agency* (Oxford: Oxford University Press).

----- (1996). *Creating the Kingdom of Ends* (New York: Cambridge University Press).

Liberty Women's Health (2016), 'Freedom to Choose',
<http://www.libertywomenshealth.com/rtc.php>, accessed 6/20/16.

Luhan, Wolfgang J.; Martin G. Kocher; Matthias Sutter (2009). 'Group Polarization in the Team Dictator Game Reconsidered', *Experimental Economics* 12 (1): 26-41

Mackie, D.M. (1986). 'Social Identification Effects in Group Polarization', *Journal of Personality and Social Psychology* 50: 720-728.

Mackie, J.L. (1977). *Ethics: Inventing Right and Wrong* (London: Penguin UK).

- Marquis, D. (2007). 'An Argument Abortion is Wrong', in Russ Shafer-Landau (ed.), *Ethical Theory: An Anthology* (Blackwell Publishing): 439-50.
- McGarty, Craig; John C. Turner, Michael A., Barbara David; et al. (1992). 'Group polarization as conformity to the prototypical group member', *British Journal of Social Psychology* 31: 1-19.
- Morini, M. (2017). 'Same-Sex Marriage and Other Moral Taboos: Cultural Acceptances, Change in American Public Opinion and the Evidence from the Opinion Polls.' *European journal of American studies*, 11(11-3).
- Morris, J.S. (2007). 'Slanted Objectivity? Perceived Media Bias, Cable News Exposure, and Political Attitudes', *Social Science Quarterly* 88(3): 707-28.
- Muldoon, Ryan (2016). *Social Contract Theory for a Diverse World: Beyond Tolerance*. Routledge.
- Myers, D.G.; S.J. Arenson (1972). 'Enhancement of dominant risk tendencies in group discussion', *Psychological Reports* 30: 615-623.
- Myers, David G.; G.D. Bishop (1970). 'Discussion effects on racial attitude', *Science* 169 (3947): 778-779

Nadeau, R., Cloutier, E., & Guay, J. H. (1993). 'New evidence about the existence of a bandwagon effect in the opinion formation process.' *International Political Science Review*, 14(2), 203-213.

Newman, L.S. (2002). *Understanding Genocide: The Social Psychology of the Holocaust* (New York, NY: Oxford University Press): 143–161.

NRLC (2016). 'Mission Statement', www.nrlc.org/about/mission, accessed 6/20/16.

Oswald, Margit E.; Grosjean, Stefan (2004). 'Confirmation Bias', In Pohl, Rüdiger F. *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory* (Hove, UK: Psychology Press): 79–96.

Parfit, D. (2011). *On What Matters*, Vols. 1&2 (Oxford: Oxford University Press).

Pew Research Center (2016). 'A Wider Ideological Gap Between More and Less Educated Adults', <http://www.people-press.org/2016/04/26/a-wider-ideological-gap-between-more-and-less-educated-adults/>, accessed on June 6, 2017.

Plous, Scott (1993). *The Psychology of Judgment and Decision Making* (McGraw Hill).

PoliticsUSA (2017). 'Proof of the GOP War on Women', <http://www.politicususa.com/proof-war-women-2>, accessed 6/5/2017.

Pojman, L.P. (1998). 'Abortion: A Defense of the Personhood Argument', *The Abortion Controversy 25 Years After Roe v. Wade: A Reader* (Cengage).

Pro-Life Perspective (2016). 'About the show', *Pro-Life Perspective*, www.prolifeperspective.com, accessed 6/20/16.

Pruitt, D. (1971). "Choice shifts in group discussion: An introductory review', *Journal of Personality and Social Psychology* 20 (3): 339–360.

Rapp, C. (2016). 'Moral opinion polarization and the erosion of trust', *Social science research*, 58, 34-45.

RealClearPolitics (2017). 'Direction of the Country', www.realclearpolitics.com/epolls/other/direction_of_country-902.html, accessed August 24, 2017.

Ross, W.D. (1930). *The Right and the Good* (Oxford: Oxford University Press, 2002).

Rorty, R. (1991). *Objectivity, Relativism, and Truth: Philosophical Papers, Vol. 1* (Cambridge: Cambridge University Press).

Rostenberg, G. (2014). 'Why Abortion is Genocide', *Live Action News*, <https://www.liveaction.org/news/why-abortion-is-genocide/>, accessed 6/2/2017.

Rowe, R. (n.d.). 'Here's Why Gun Control Doesn't Make Sense Right Now.' *Ranker*, <https://www.ranker.com/list/arguments-against-gun-control/richard-rowe>, accessed 12/28/2018.

Scanlon, T.M. (1998). *What We Owe to Each Other* (Cambridge, MA: Harvard University Press).

Shafer-Landau, R. (2003). *Moral Realism: A Defense* (Oxford: Oxford University Press).

Sheehan, Cindy (2006). *Peace Mom: A Mother's Journey Through Heartache to Activism* (Atria Books).

Sheikh, H., Ginges, J., Coman, A., & Atran, S. (2012). 'Religion, group threat and sacred values.' *Judgment and Decision Making*, 7(2), 110.

Shoemaker, D.W. (2011). 'Psychopathy, Responsibility, and the Moral/Conventional Distinction', *Southern Journal of Philosophy*, 49(s1): 99-124

- Stoner, J.A. (1961). 'A comparison of individual and group decision involving risk', *Unpublished master's thesis, Massachusetts Institute of Technology*.
- Stratton-Lake, P. (2014). 'Intuitionism in Ethics', in E.N. Zalta (ed), *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition), <http://plato.stanford.edu/archives/win2014/entries/intuitionism-ethics/>, accessed 6/16/2016.
- Strawson, P.F. (1962). 'Freedom and resentment', *Proceedings of the British Academy* 48:1-25.
- Stroud, N.J. (2010). 'Polarization and partisan selective exposure', *Journal of Communication*, 60(3), 556-576.
- Sunstein, C.R. (2002). 'The Law of Group Polarization', *The Journal of Political Philosophy* 10 (2): 175-195
- Tawia, S. (1992). 'When is the capacity for sentience acquired during human fetal development?', *Journal of Maternal-Fetal Medicine*, 1(3), 153-165.
- Thomson, J.J. (1976). 'A Defense of Abortion', in *Biomedical Ethics and the Law* (Springer).
- Timmons, M. (2007). *Disputed Moral Issues: A Reader* (New York: Oxford University Press).

- Tosi, J. & Warmke, B. (2016). 'Moral Grandstanding', *Philosophy and Public Affairs* 44(3): 197-217.
- Van Roojen, M. (2015). 'Moral Cognitivism vs. Non-Cognitivism', in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition), <http://plato.stanford.edu/archives/fall2015/entries/moral-cognitivism>, accessed 6/20/16.
- Vaughn, L. (2009). *Bioethics: Principles, Issues, and Cases* (Oxford: Oxford University Press).
- Vinokur, A. & Burnstein, E. (1974). 'Effects of partially shared persuasive arguments on group induced shifts: A group problem-solving approach', *Journal of Personality and Social Psychology* 29: 305–315.
- Walker, T.G. & Main, E.C. (1973). 'Choice shifts and extreme behavior: Judicial review in the federal courts', *The Journal of Social Psychology*. 2 91 (2): 215–221.
- Warren, M.A. (1973). 'On the Moral and Legal Status of Abortion', *The Monist*: 43-61.
- Yardi, S. & Boyd D. (2010). 'Dynamic Debates: An analysis of group polarization over time on Twitter', *Bulletin of Science, Technology and Society* 30 (5): 316–27.

Zhang, Y., Lewis, M., Pellon, M., & Coleman, P. (2007). 'A Preliminary Research on Modeling Cognitive Agents for Social Environments in Multi-Agent Systems.' In *AAAI 2007 Fall Symposium, Emergent Agents and Socialities: Social and Organizational Aspects of Intelligence*: 116-123.