

Naturalistic Moral Realism and the Recent Empirical Findings from Social Psychology

Ryo Chonabayashi

1. Introduction

Many contemporary philosophers often appeal to empirical findings and either defend their positions or raise objections to their rival views. The recent wave of so called ‘experimental philosophy’ is an instance of this state of affairs. For instance, some philosophers have attempted to shed light on some old philosophical issues such as free will, determinism and moral responsibility, by appealing to the empirical findings of lay people’s judgements about action in certain context. Other philosophers have attempted to provide some new insights concerning the nature of knowledge, which is another old philosophical issue, by appealing to the empirical findings of people’s judgements about when we know things and when we do not. What these examples indicate is that some philosophical questions may be answered if we look at the empirical data studied and discovered in the relevant sciences.

The idea that philosophy can advance with the help of science is not new. Historically, John Stuart Mill claimed that

we can gain knowledge about morality by ‘observation and experience’, and what Mill’s claim indicates is the idea that we can advance our knowledge about morality by appealing to empirical science which is based on our experience (Mill 1871/2002).

This epistemological thesis about moral knowledge is not just an old philosophical doctrine. The thesis has some contemporary supporters, and one group of philosophers who attempt to defend this thesis is called the ‘Cornell realists’ (Sturgeon 1985/1998, Boyd 1988, Brink 1989). The name is given due to their association with Cornell University.

The Cornell realists hold not just this epistemological thesis that but also hold related claims concerning morality. For example, they hold that there are so called ‘moral properties’, such as wrongness manifested in torturing innocent children, rightness exemplified in distributing goods fairly and courage instantiated in someone’s character. They also hold these moral properties are ‘natural properties’ and they are part of the natural world in the same way various properties studied in the empirical sciences are. They also claim that the definitions of moral terms can be given in the same way we can give the *a posteriori* definitions of natural kind terms. Finally, they claim that first-order moral theories are supposed to provide such *a posteriori* definitions of moral terms. In this essay, let us use the term ‘naturalistic moral realism’ to refer to a set of these claims about morality.

At first glance, this sort of naturalistic position in moral philosophy seems to go hand in hand with the recent trend in

philosophy. The central idea of the recent trend is that we can shed light on some philosophical issues by appealing to the relevant sciences. On the other hand, what naturalistic moral realism says is that morality is a sort of science and our best normative theory can provide us with *a posteriori* definitions of moral terms in the same way our best theories of physics provide us with *a posteriori* definitions of subatomic particles. If our moral knowledge is *a posteriori*, the way we obtain such knowledge must be based on various empirical data. So, appealing to relevant scientific theories should be crucial for gaining such *a posteriori* moral knowledge.

Despite this expectation, there is tension between the recent empirical trend in philosophy and this naturalistic position in moral philosophy. The worry is that some empirical findings about moral judgements might in conflict with one important argument which is supposed to defend naturalistic moral realism.

In this essay, I shall discuss this possible tension and consider how naturalists can reply to this worry. The structure of this essay is as follows. First, I shall give a concise account of naturalistic moral realism and how this position may be undermined by the empirical findings discovered in social psychology. Then, I shall propose three replies to the idea that the empirical findings from recent social psychology undermines naturalistic moral realism.

2. Naturalistic Moral Realism and the Theory Argument

Naturalistic moral realism can be more formally characterised as follows:

(C) There are mind-independent natural moral properties whose instantiation results in moral facts, and these moral properties are investigated by empirical ways of reasoning.

The defence of naturalistic moral realism amounts to the defence of this claim.

(C) contains a *moral realist thesis* according to which there are mind-independent moral properties. The mind-independence of moral properties can be understood as the *stance-independence* of them: moral facts are stance-independent in the sense that they are not constituted by the ratification from any actual or hypothetical perspective (Shafer-Landau 2003). For instance, the fact that killing is wrong is not so because the prohibition of killing others is agreed in any actual or hypothetical community. The fact obtains regardless of whether any particular actual or hypothetical community agree or disagree on the norm which prescribes.

Also, notice I use the term ‘moral properties’ in (C). Thus, I take that moral entities such as moral goodness, wrongness, justice, courage, etc., can be understood as moral properties.

(C) also contains the thesis that moral properties are natural properties. A rough explanation of this thesis is that moral properties are ‘natural properties of the same general sort

as properties investigated by the sciences' (Sturgeon 2005, p. 92). This characterisation of moral properties roughly explains how moral properties are part of the natural world. The natural world where we human beings are living is the object of scientific inquiries, and properties investigated in science are instantiated in the objects in this natural world. The property of being negatively charged is a property of an elementary particle, which is, in turn, part of the natural world. If moral properties are of the same general sort as properties investigated by the sciences, moral properties are also part of the world in the same way other properties investigated in science are.

In the contemporary literature on naturalistic moral realism, one particular argument for this position has been lively discussed and that argument has been an object of serious scholarship in metaethics. Here is the argument:

The Explanationist Argument for Naturalistic Moral Realism

- (1) An entity exists if that entity is ineliminable in the best explanation of phenomena we experience.
- (2) Natural moral properties are ineliminable in the best explanation of phenomena we experience.
- (3) Therefore: natural moral properties exist.

Since this argument appeals to the explanatory virtue of moral properties, it is sometimes called 'the explanationist argument' (Sinclair 2011).

One of the underlying thoughts of this argument is an analogy between science and ethics. In science, a theory

which postulates theoretical entities such as scientific laws may be confirmed due to its *explanatory virtue*: we believe in the laws of nature posited in the theory because we can explain empirical phenomena by postulating the existence of those laws. The proponents try to defend the claim that *moral properties have similar explanatory virtue*: there are some empirical phenomena which are explained in terms of natural moral properties.

The central issue concerning this argument is about the second premise. At first glance, it is not clear how the assumption that there are moral properties contribute to our explanatory practice can be defended. Certainly, moral properties seem to be explanatorily relevant in some way. We often try to give an explanation of why Taro's hitting his little brother was wrong by appealing to some non-moral facts about Taro's act. Perhaps his act was wrong due to the fact that Taro had a clear intention to harm his little brother. The explanation sought in this context employs the notion of 'wrong'. However, this case is not helpful for defending the second premise of the explanationist argument. What this case indicates is that the assumption of moral properties may be explained by some non-moral facts. But what is needed for the second premise are cases in which the assumption of moral properties explains some non-moral facts. Then, how can the assumption of moral properties, such as the assumption concerning the wrongness of Taro's act, explain non-moral facts?

In the contemporary literature, it is thought that one way to defend the second premise is to offer so called, 'moral

explanations'. This line of argument is originally suggested by Nicholas Sturgeon and his suggestion stimulated lively discussions (1985/1998, 2006). I call this line of argument 'the singular moral explanationist argument' (the singular argument, for short).

Sturgeon suggests several types of moral explanations. For instance, Sturgeon gives some cases in which moral facts about character traits explain the formation of moral beliefs. One such case is this: Hitler's character trait explains why we have the moral belief that Hitler was morally depraved. Another type of moral explanation Sturgeon uses is cases in which moral facts explain certain historical facts. Sturgeon suggests that the fact that chattel slavery in a few countries was much worse than previous forms of slavery explains why vigorous and reasonably widespread moral opposition to slavery arose for the first time in the eighteenth and nineteenth centuries primarily in Britain, France, and English speaking North America (cf. Miller 1985, p. 527).

Sturgeon's proposal implies that natural moral properties are ineliminable in the best explanation of phenomena we experience, so the success of the singular argument may support the second premise of the explanationist argument.

It seems to me, though, there is another way to defend the second premise. I call this second way 'the first-order theory argument for moral realism' (the theory argument, for short). Boyd presented this line of argument in his influential paper, 'How to be a Moral Realist' (1988). In his paper, Boyd suggests that an argument for moral realism which is analogous to one

influential argument for scientific realism is possible.

In order to see the argument for moral realism Boyd suggests, let us see the argument for scientific realism first.

Argument for Scientific Realism

- (1) Scientific theories are empirically reliable.
- (2) The theory-building procedures of science are not free from theoretical presuppositions.
- (3) Therefore: scientific realism is true.

The warrant for the conclusion is that scientific realism best explains the empirical success of scientific theories produced by theory-dependent methods. This argument appeals to empirical facts about science [(1) and (2)], and provides a scientific realist explanation as a scientific hypothesis (Boyd 2002).

What Boyd suggests is that we can have an argument for moral realism which is analogous to this argument for scientific realism. Here is the argument.

The First-Order Theory Argument

- (1) First-order ethical theory is empirically reliable.
- (2) The theory-building procedures of first-order ethical theory are not free from theoretical presuppositions.
- (3) The best explanation of (1) and (2) [which are themselves empirical phenomena] is moral realism which implies the existence of natural moral properties.
- (4) Therefore: natural moral properties are ineliminable in the

best explanation of phenomena we experience [the second premise of the explanationist argument].

Unlike the singular argument, this argument does not appeal to the explanatory virtues of singular moral explanations. Rather, it appeals to the explanatory virtues of the hypothesis that moral realism is true. Also, the theory argument appeals to some features of substantial first-order ethical theories. Although the singular argument also appeals to some substantial ethical claims such as that Hitler was morally depraved, it does not refer to the features of ethical theories which may be the source of such particular moral claims.

Although it is apparent that these two ways to defend the second premise of the explanationist argument are different in these ways, it seems that the distinction between them is not properly recognised in the literature. That is partly because the theory argument has not been clearly articulated and has not been differentiated from the singular argument. Boyd's essay is often mentioned as a paper which supports the explanationist argument, but main discussions of the papers which mention Boyd's essay tend to focus on the singular argument (cf. Darwall et al. 1992; Morgan 2006; Rea 2006). Consequently, there are not many discussions on the theory argument¹.

It is not true that Boyd's paper has been unfairly neglected. The situation is opposite: Boyd's paper stimulated lively discussions in contemporary metaethics. Boyd's essay has been discussed by many philosophers due to the fact that the externalist semantic theory proposed in his essay has

been regarded as an important suggestion for moral realism. Nevertheless, there is relatively little discussion on the theory argument Boyd suggests which is also crucial part of the full defence of naturalistic moral realism.

At the moment, there are some objections to the singular argument. Some argue that moral properties are not causal entities and it is not appropriate to ascribe causal power to them (Harman 1986, p. 63; Thomson 1996). Another objection is that moral explanations are not the best explanation and there are non-moral explanations which better explain the empirical phenomena Sturgeon appeals to (Leiter 2001). The proponents of this objection argue that people's moral belief about Hitler is best explained by some non-moral facts, not by the moral fact that Hitler's character was evil. This is a simple but important objection since many people might be inclined to explain the phenomena Sturgeon uses in terms of facts about human psychology and evolution, rather than the existence of moral properties.

It might be possible to provide some persuasive replies to these objections against the singular argument, but given the fact that these objections have been somehow influential, it would be good if the proponent of naturalistic moral realism got another argument for his position. This is the theory argument is an important one for the defence of naturalistic moral realism.

3. Social Psychology Undermining the Theory Argument

In the previous section, I explicated the central claims of naturalistic moral realism and presented two arguments for it. I also provided some reason for considering the theory argument seriously in order to defend naturalistic moral realism. In this section, I shall show that one of the premises of the theory argument may be in conflict with some recent empirical findings from social psychology.

The second premise of the theory argument implies that first-order ethical theory is developed in the same way science is developed. In the revision procedures of science, the following theory-dependent features are typically seen: (1) hypotheses presuppose background assumptions and theories, (2) central issues depend on the theoretical context, (3) the standard of confirmation depends on background theories, and (4) there is the mutual relation between old theories and revised theories. The second premise implies that moral inquiry also possesses these theory-dependent features.

The second premise is an empirical premise. Thus, the premise is rejected if empirical research shows that first-order ethical theory does not have such theory-dependent features.

These theory-dependent features presuppose some sort of rational theorising. When theorists suggest new hypotheses referring to theoretical entities, they suggest new hypotheses on some reasons. Through some sort of rational reflection on background assumptions, theorists suggest new hypotheses. In such reflection, theorists might see the importance of new

hypotheses. Theorists might find the coherence between new hypotheses and background theories and assumptions. Theorists do these things by recognising some reasons for their theorising.

Some recent studies in social psychology have been held to raise doubts about the role of rational theorising in morality. They seem to be showing that there is no genuine link between the formation of moral judgements and moral reasoning as traditionally assumed by philosophers. Further, these studies indicate that how people form moral judgements is radically different from how scientific theories are built up. Such empirical results seem to be undermining the second premise of the theory argument: if the way we form moral judgements was radically different from the way scientists build up their theories, moral inquiry would not have those theory-dependent features analogous to the theory-dependent features of science. Hence, the second premise of the theory argument is empirically rejected. Below, I shall explain this worry in detail.

In social psychology, the mechanism of how moral judgements are formed is an important subject. The rationalist model and the social intuitionist model are two competing models of how we form moral judgements. Jonathan Haidt argues that recent empirical studies show that the rationalist model does not correctly describe how moral judgements are formed (2001). Haidt argues that the alternative social intuitionist model more accurately describes the mechanism of moral judgements.

Consider the following moral judgement, ‘it is wrong for

Tom to hit his innocent little brother for merely making fun'. According to the rationalist model, this moral judgement is formed mainly by the process of reasoning and reflection. In making the judgement, the judge might be considering the following aspects of the act: what sort of effect, whether some good effects and bad effects would come out of such an act, and whether there could be any case where an innocent person's being hit is justifiable. Through considering various morally relevant aspects of the act, the judge makes the judgement that the act is morally wrong. Reasoning and reflection play important roles in this judgement formation process. By reasoning and reflection, the judge carefully analyses various aspects of the act and reaches the moral conclusion.

According to the social intuitionist model, on the other hand, the formation of this moral judgement starts with initial reactions to the case given. First, the judge 'feels' that hitting an innocent person for merely making fun is morally wrong (Haidt 2001, p. 814). Then, when the judge faces with a social demand for a verbal justification, the judge becomes a 'lawyer' trying to build a case rather than a judge searching for the truth (ibid). If the judge is asked why the judge believes that hitting an innocent person for fun is morally wrong, the judge might provide some reasons for the claim. But, according to the social intuitionist model, the consideration of these reasons does not play an essential role in the formation of the judge's belief. Rather, these reasons are provided just to defend the judge's initial reaction that hitting an innocent person for fun is wrong. So, the role of reasoning and reflection is different from the

role played in the rationalist model. Reasoning and reflection are secondary in this judgement formation process. Rather than searching truth by reasoning and reflection, the judger employs reasoning and reflection to defend the judger's initial reaction to the case.

Haidt writes that the difference between the rationalist model and the social intuitionist model can be metaphorically described as the difference between a lawyer defending a client and a scientist seeking truth (Haidt 2001, p. 820). The former is analogous to the social intuitionist approach and the latter is analogous to the rationalist approach. While the rationalist model says that the judger tries to find an objective answer to a given case like a scientist, the social intuitionist model says that the judger tries to defend his initial reaction like a lawyer who tries to defend his client.

Haidt gives a series of empirical findings as evidence for the social intuitionist model.

Certain Social Settings Significantly Affect our Moral Judgements

The studies on attitudes, people's perception and persuasion show that desires for keeping harmony and agreement have significant effects on our judgements (Haidt 2001, p. 821). For instance, one study shows that one's initial attitudes toward some controversial issues will be shifted toward one's partner's view if one knows the partner's view prior to the discussion (Chen et al. 1996). Another study shows that if people are expected to work with a particular person, their judgement

about this person would be more friendly than the case when people are not expected to work with that person (Darley and Berscheid 1967). These are the instances of how our desires to keep harmony and agreement in our environments affect our judgements.

These findings indicate that our moral judgements are also significantly influenced by certain social factors. My judgement about Rob who is one of my friends might be different if Rob is a complete stranger. My judgement about the capital punishment might be affected by my discussion partner's view if I am about to engage a dialogue on this issue with a person who supports the practice. In this way, the formation of our moral judgements is highly influenced by social settings, and reasoning and reflection alone are not the main factors which form moral judgements.

Defence Motivation

Another study shows that people have a desire called, 'defence desire'. This is people's desire to hold attitudes and beliefs that are congruent with existing self-definitional attitudes, which includes moral commitments and beliefs (Chaiken et al. 1996). When defence motivation is triggered, both heuristic (intuitive system) and systematic thinking (reason system) work to preserve self-definitional attitudes (rather than seeking truth). One study shows that when students are asked to research evidence on both the plausibility and implausibility of the death penalty, students accept evidence which supports their prior belief uncritically while they carefully interpret opposing

evidence (Lord and Ross 1979).

This study also shows that the role of reasoning and reflection in moral judgements is to defend prior moral commitments. The role of them is not to find the true answer (Haidt 2001, p. 821).

Cognitive Dissonance

Although Haidt does not mention this study in his article, another relevant phenomenon is observed in the field about cognitive dissonance. According to cognitive dissonance theory, when people have two contradicting cognitions, such as two beliefs whose truths are incompatible, in order to reduce their unpleasant psychological state which is generated by the dissonance of these competing cognitions, people tend to (1) change their cognitions, (2) add more consonant cognitions, or (3) change their views on the values of these cognitions (cf. Cooper 2007). This may be applied into the case of moral judgement. If people have two competing moral beliefs, they will be in a psychologically unpleasant state, and in turn, try to reduce this unpleasantness by changing their moral beliefs. This may be taken as further empirical evidence for the social intuitionist model: people change their moral views in order to reduce their psychological unpleasantness, not because they find some theoretical reasons for the revision.

Ad-Hoc Moral Reasoning

The studies on how children develop their moral views also support the social intuitionist model. Kohlberg's theory used to

suggest that young children in many cultures first hold the view that acts that get punished are wrong and acts that get rewarded are good. Then, they soon advance to the theory that acts that others approve of are good while acts condemned by others are bad (Kohlberg 1969). But more recent studies show that those children who are supposed to be at the stage where they believe that acts that get punished are wrong actually believe that harmful acts are wrong even if those acts are not punished (Turiel 1983). According to Haidt, this later finding shows that children form their moral judgements by their automatic moral intuitions. The reasons those children could provide for the justification of their judgement, such as, ‘I thought the act was wrong since it would be punished’, is *ad hoc* justification which just ‘sounds’ plausible (Haidt 2001, p. 823).

Given these empirical findings, Haidt concludes that the social intuitionist model is more empirically plausible than the rationalist model. He writes:

‘[These empirical findings are] intended to demonstrate that the roots of human intelligence, rationality, and ethical sophistication should not be sought in our ability to search for and evaluate evidence in an open and unbiased way’ (Haidt 2001, pp. 821-822).

If...shocking or threatening issues are being judged, such as abortion, euthanasia, or consensual incest, then coherence motives [the desire to keep the existing moral commitments] also will be at work. Under these more realistic circumstances, moral reasoning is not left free to search for truth but is likely to be hired out like a lawyer by

various motives, employed only to seek confirmation of preordained conclusions' (Haidt 2001, p. 822).

Suppose that Haidt's interpretation of these empirical findings is correct and the social intuitionist model is an accurate explanation of how we make moral judgements. What does this conclusion show? How would the social intuitionist model be a threat to the theory argument?

The social intuitionist model might be a threat to the theory argument since this model might undermine the second premise of the argument. As we have seen in the beginning of this section, the second premise of the argument says that moral inquiry is analogous to scientific enquiry in the way it is conducted on certain reasons. If the social intuitionist model is true, moral inquiry should be understood as an inquiry which is not to seek truth, rather to defend our initial reactions to given cases.

Remember the four theory-dependent features of science: (1) hypotheses presuppose background assumptions and theories, (2) central issues depend on the theoretical context, (3) the standard of confirmation depends on background theories, and (4) there is the mutual relation between old theories and revised theories. If the social intuitionist model was true, moral inquiry would not have these theoretical features.

First, the truth of the social intuitionist model seems to be in conflict with (1). According to the social intuitionist model, moral hypotheses we make are not the product of reasoning or reflection. For a hypothesis to presuppose background

assumptions and theories, such reasoning or reflection prior to the suggestion of that hypothesis is needed. The judger needs to consider a hypothesis from the candidates which seem to be not in conflict with background moral assumptions. Such a procedure requires that the judger proposes a hypothesis on the reflection about the coherence between the hypothesis and background assumptions. But, the social intuitionist model says, such consideration is not essential to the formation of moral judgements. Moral hypotheses are suggested not by presupposing certain background moral assumptions. Rather, they are suggested by the judger's initial reactions to given cases.

Because of a similar reason, the social intuitionist model is in conflict with (2). Since moral inquiry is primarily concerned with initial moral judgements, there is no genuine room for the judger to consider what sort of thing would be issues inherited from the past theory. We just feel that such and such things are morally important, rather than giving answers to certain questions inherited in the theoretical tradition.

These considerations may support the thought that the truth of the social intuitionist model of moral judgements is a threat to the theory argument. From these, the opponent of the theory argument would argue as follows: the empirical findings support the social intuitionist model, and the social intuitionist model rejects the rational aspects of moral inquiry which is needed for the defence of the second premise of the theory argument. So the theory argument should be rejected, the opponent would conclude.

4. Lawyers Defend her Case on Reasonsⁱⁱ

I have three replies to this objection against the theory argument. The first reply goes as follows: even if the social intuitionist model is true, the truth of the social intuitionist model does not become a real threat to the theory argument. Even if the main factor of moral judgements is our initial judgements, our judgements and the defence of them must be on some reasons. If this is the case, we could defend the theory argument from the threat of the empirical objection.

Consider the following hypothetical case. Linda, a lawyer, attempts to defend her client who is trying to come into the inheritance of his parents. There is a quarrel over this inheritance because other members of family are also trying to come into the inheritance. For Linda, it does not matter if the client really deserves it. Her job is to give as many points which attract the court as possible. In this case, Linda's basic position, 'my client deserves getting the inheritance', does not change though there is a real legal possibility that the client is not entitled to come into the inheritance.

Now, consider the following case of Mario who makes moral judgements. He considers whether killing an innocent person for the killer's fun is permissible or impermissible. Mario believes that such an act is not permissible. He could provide with us some reasons for his belief. But, the empirical findings appealed by the social intuitionist show that his basic position is not from his rational reflection of the case. Rather,

he has the initial feeling that such an act is not permissible, and on the basis of this initial reaction he forms the belief. His providing some reasons for his belief is *ad hoc* in the sense that his basic belief remains the same and all the reasons he would refer to are just to defend his basic position. This is analogous to the way Linda attempts to defend her client.

Now, let us see Linda's case again. As the scenario mentions, Linda is capable of giving some reasons for her case though her basic position remains the same. But, how does she give such reasons for her case? It seems that she has to refer to some reasons which could be accepted by other people even if their position is different from hers. Otherwise, she cannot persuade the judges. The way she refers to such reasons is, it seems, theory-dependent in the same way science is. She should refer to some facts which seem to be defending the client from the perspective of the confirmation standard inherited in her legal tradition. She might mention the fact that her client was supporting his parents when they were still alive since, from the confirmation standard, this sort of fact could be regarded as evidence for the client's right. Also, She should refer to some facts which seem to be weakening her opponent's case. She might mention the fact that other members of the client's family did not receive any documental proof which indicates the parents' will. She would mention such a fact since she knows that this sort of fact would be the central issue in the court. She could even revise her position if some unknown facts which hugely undermine her case suddenly appeared. Her opponent might show us some evidence which indicates that the parents

in fact hoped that their family equally shared the inheritance. Given this new evidence, she might revise her position in the following way: ‘now the new evidence shows that the parents had the will that the inheritance should be distributed equally. But, this should not undermine the client’s right to come into large part of the inheritance since he is legally entitled to such a claim’. All of these considerations seem to be analogous to the theory-dependent features of science.

Now, analogously, let us see Mario’s case again. In this scenario, we are assuming that the social intuitionist model is true and Mario’s moral belief is formed primarily by his initial reaction to the case. Now, suppose Mario has a friend, Saif, who is sceptical about morality. There is a quarrel over whether such an act is acceptable or not. Saif says that such an act is in the end permissible since he believes that there is no objectivity in morality. Saif asks Mario to provide some reasons for Mario’s claim. Mario could provide some *ad hoc* reasons which are just to defend Mario’s position, rather than for seeking the truth of the proposition in question. He has to refer to some reasons which could be accepted by Saif. Otherwise, Mario cannot change Saif’s mind. The way he refers to such reasons is, again, theory dependent in the same way science is. He should refer to some facts which seem to be defending his position from some basic understandings of morality. He might mention the fact that such an act would affect the sufferings of the victim, the victim’s family and friends, since this sort of fact seems to be morally relevant from the perspective of such basic understandings of morality. He might mention that

Saif's holding a sceptical position would undermine not only the foundation of morality, but also the foundations of various parts of our society. He would mention such a fact since he thinks that this sort of fact would be one of the central issues in the debate. He could even revise his position if Saif raised some points which potentially undermine his position. Saif might argue that we could not touch and see value and there must be no such moral value, and Mario is committed to the existence of an entity which we do not have any reason to believe it. Given this objection, Mario might revise his belief in the following way: 'now, it is difficult to say that killing an innocent person for fun is impermissible since we cannot see or touch the impermissibility of such an act. But, we have various other reasons to believe that such an act is impermissible since the acceptance of the social code of such an act would produce various sufferings in the world'. For Mario, Saif's objection is rather a springboard to strengthen his view.

All of these considerations seem to be analogous to the theory-dependent features of science. In defending his hypothesis, Mario refers to some background moral assumptions and this seems to be analogous to the way scientists suggest a new hypothesis: they suggest a hypothesis on the basis of background theories they accept. Mario's disagreeing with Saif itself may be seen as an aspect of ethical inquiry analogous to the way science is conducted. This judgement (Mario's disagreement with Saif is morally significant) is dependent on his background conception of morality, such as that the wrongness of killing is one of

the central issues in ethics. This may be seen as analogous to scientists' judgement that some particular theoretical issues are important and need to be pursued further. Mario's sophistication of his hypothesis facing Saif's challenge may be seen as analogous to the mutual relation between old theories and revised theories in science. Mario's sophisticated response is a revision of his initial reaction, and we can see the mutual relation between his initial reaction, Saif's objection and the sophisticated answer.

The upshot is that even if the way moral judgements are formed is analogous to the way a lawyer defends her case, we could still see some important theory-dependent features in both forms of inquiry which are analogous to the way science is developed. Therefore, we can accept the empirical findings on which the social intuitionist model is established without giving up the idea that moral inquiry is analogous to scientific inquiry.

5. Lay People and Expert Moral Theorists

There is another reply to the empirical objection against the theory argument. This reply relies on the distinction between lay people and expert moral theorists. The reply goes as follows:

- (1) To undermine the second premise of the theory argument, it needs to be shown that moral inquiry conducted by expert moral theorists is not analogous to the way science is developed.

- (2) The recent studies in social psychology merely show that moral inquiry conducted by lay people is not analogous to the way science is developed.
- (3) Therefore: the recent studies in social psychology do not undermine the second premise of the theory argument.

About (1): to undermine the theory argument, empirical research needs to show the psychology of moral theorists who conduct serious theory-construction of morality. That is because the second premise of the theory argument is not about the psychology of lay people. Rather the second premise is about how moral theorists build up their theories. This point can be clear when we remind ourselves the thought that the theory argument is analogous to the abductive argument for scientific realism. In the argument for scientific realism, the premise about the theory-dependent features of science is not about how lay people form their beliefs about physics. It does not matter how lay people form their beliefs about subatomic particles by reading physics textbooks, or by putting some initial thoughts on the matter. What matters is how real scientists build up their theories by setting up experiments, observing the results of these experiments and interpreting these observations. So, analogously, what matters is how real moral theorists build up their normative theories, rather than how lay people form their moral judgements.

About (2): the empirical findings Haidt refers to are all about the psychology of lay people. Haidt himself explicitly says that the purpose of his research is to reveal the process of

forming moral beliefs in real life setting, not to find out how people rationally construct their moral theory (Haidt 2001, p. 820). So, the premise (2) is not implausible.

Interestingly, some of the empirical studies Haidt uses might support the claim that the way expert moral theorists conduct their normative theories is analogous to the way science is developed. Some of the recent empirical studies Haidt uses show that when people are in appropriate circumstances and asked to construct their impartial moral theories, they can construct their moral theories mainly by reasoning and reflection. According to Haidt, some empirical studies show that when a person is under certain circumstances (e.g. a person is given adequate time to consider cases, a person is in a position where that person's judgements are not about someone to whom that person knows or has some special relationships, etc.) biased motivations are not triggered (Haidt, p. 822). This could support the thought that moral theorists who are supposed to be in such circumstances should be able to conduct their theory-building mainly by reasoning and reflection.

Since (1) and (2) are plausible, (3) should also follow from these two premises: the recent studies in social psychology Haidt uses are not a real threat to the second premise of the theory argument.

These considerations make it clear what sort of empirical findings could be a threat to the second premise of the theory argument. A sort of empirical findings which could be a threat to the theory argument is empirical studies on the psychology

of moral theorists who build up normative ethical theories.

Are there such empirical studies which focus on the psychology of normative ethicists? There are some empirical studies on how normative ethicists behave (cf. Schwitzgebel 2009; Schwitzgebel and Rust 2009). But these are not the ones we are seeking because we are seeking certain empirical studies on how normative ethicists build up their theories. More relevant studies may be found in the work of sociologists who work on the historical research on the development of moral philosophy (Collins 2000). Such studies might show that in the development of moral philosophy there are certain theoretical features analogous to the theory dependent features seen in science. Or, they might undermine the analogy between the theory building procedures in ethics and scienceⁱⁱ. In any case, the empirical findings provided by the recent studies from social psychology are not the sort of the things which undermine the theory argument.

6. How Normative Ethicists Build Up Their Theories

My third reply amounts to the claim that we could in fact show that there are some analogous features between the theory building procedures in ethics and science by seeing how contemporary normative ethicists build up their theories. The spirit of this response is that how real normative ethicists develop their theories *in fact confirms* the second premise of the theory argument. Below I describe how Brad Hooker, who is an influential contemporary normative ethicist, develops his

theory. Then, I examine whether the way Hooker develops his theory is analogous to the way science is developed.

In his *Ideal Code, Real World*, Hooker defends his version of rule consequentialism. Hooker's consequentialism seeks the ideal moral code whose acceptance is reasonably expected to produce more aggregate value than any other code. The acceptance of a code means its collective internalisation. For instance, the internalisation of the moral code of keeping promises amounts to the establishment of a shared conscience in society where people believe or feel^{iv} that they should keep promises (Hooker 2000, pp. 2, 5). From this, Hooker gives an account of a wrong act: an act is wrong if everyone's feeling free to do it would have bad consequences (Hooker 2000, p. 5).

How do we seek the ideal moral code? Hooker suggests the following five criteria for the assessment of moral theories.

- (1) Moral theories must start from attractive general beliefs about morality.
- (2) Moral theories must be internally consistent.
- (3) Moral theories must cohere with the moral convictions we have after careful reflection.
- (4) Moral theories should identify a fundamental principle that both (a) explains why our more specific considered moral convictions are correct and (b) justifies them from an impartial point of view.
- (5) Moral theories should help us deal with moral questions about which we are not confident, or do not agree (Hooker 2000, p. 4).

Hooker's theory-building starts with (1). He suggests following moral convictions as attractive general moral beliefs.

Attractive General Moral Beliefs

- (a) We ought to help others in need, even those with whom we have not special relationships, unless helping those in need does not require great sacrifice of ourselves (Hooker 2000, p. 14).
- (b) We owe more altruism to certain people, such as families and friends, than others.
- (c) There are certain acts, such as physically attacking innocent people and their property, taking others' property without having any agreement, lying, and breaking promises, which are morally impermissible though in certain circumstances those acts may be permissible or required (Hooker 2000, p. 17).

Then, Hooker goes on to the procedure (4); seeking the moral principle which explains why those moral beliefs are plausible, and justifies them from an impartial point of view. The justification of those moral beliefs is that people's accepting these beliefs and living in accord with the internalisation of the codes of those beliefs have the best overall consequences (Hooker 2000, p. 4). Hooker's theory is obviously impartial in the sense that we can assess the plausibility of those moral convictions by observing the consequences of the internalisation of the codes of those practices.

Now, what we want to see is whether there is any analogous feature between Hooker's theory and science in general. In the case of science, certain theory-dependent features are seen when some revisions in theories occur. So, the question we should ask is whether we can see any such theory-dependent feature in the revision process of Hooker's theory. We can see how Hooker might revise his theory in his response to one objection against consequentialism. I call the objection 'the calculation objection'. The objection goes as follows:

The Calculation Objection

- (1) Consequentialism implies that we ought to act in the way the consequences of everyone's feeling morally required to do a given act would be better than the consequences of everyone's not feeling so.
- (2) In order to know the code which should be internalised, we need to know what sort of expected value we would gain by internalising that code.
- (3) (2) is practically impossible (cf. Griffin 1996, p. 107)
- (4) Therefore: consequentialism is practically impossible.

In response to the calculation objection, Hooker suggests what he calls 'wary consequentialism'. In response to the objection, Hooker first claims that what we can find is a code 'whose general internalization could reasonably be expected to result in at least as much good as could reasonably be expected to result from any other *identifiable* code' (Hooker 2000, p. 114). Then, Hooker continues:

So suppose we instead aspire to find a code whose general internalization could reasonably be expected to produce as much good as any other code we can identify. More than one code may pass this test. That is, more than one code may have unsurpassed expected value. Rule-consequentialism must have a way of selecting among the codes in this set. Suppose rule-consequentialism is formulated so as to claim that, of these codes with unsurpassed expected value, the one closest to conventional morality determines which kinds of act are wrong. Call this view *wary rule-consequentialism* (Hooker 2000, p. 114).

One thing that counts in favour of wary rule-consequentialism is its epistemological modesty. We start with what we know, with what has been already tried. Attempts at moral reform should begin with existing practices, and then prune, refine, and supplement these where changes seem very likely to increase the overall good (Hooker 2000, p. 115).

Here we can see one criterion of the revision procedure Hooker's consequentialism has. The consequentialist theory is revised if the alternative moral code to existing practices is highly likely to increase expected value (Hooker 2000, p. 116). If there are two moral hypotheses with unsurpassed expected value suggested, the one which is closer to existing practices is chosen.

Above, I elucidated the basic structure of Hooker's theory and how he revises his theory. Are the four theory-dependent

features analogous to the ones of science seen in Hooker's theory-building procedures? Or, is the way Hooker develops his theory further evidence for the social intuitionist model and, in turn, evidence for the claim that the way ethics is developed is different from the way science is conducted? I claim that theory-dependent features analogous to the way science is conducted can be seen in Hooker's theorising. I exhibit these theory-dependent features of Hooker's theory through explaining his theory further.

Hypotheses Presuppose Background Assumptions and Theories

A new hypothesis in Hooker's theory presupposes the plausibility of basic moral convictions. This point can be seen when Hooker tries to give a rule-consequentialist answer to the question about how much the relatively well-off are obliged to do for the needy (Hooker 2000, p. 159). Hooker suggests the following moral rule: agents are required to help those in greater need, especially the worst off, even if the personal sacrifices involved in helping others add up to a significant cost to the agents over the course of their lives. Agents who accept this rule will be disposed to help those in greater need, and to do so up to at least the threshold of 'significant aggregative personal cost' (Hooker 2000, p. 174). Someone might think that the suggestion of this principle does not require us to refer to any background assumptions of morality since the principle itself is intuitively plausible. But Hooker suggests this hypothesis as a hypothesis which coheres with

two of our assumptions about morality, that we have duty to others in need and there are limits on how much self-sacrifice morality can reasonably demand. This is how Hooker follows the five criteria for the assessment of moral theory. Here we can see that Hooker is suggesting the hypothesis since he sees his hypothesis is a reasonable one from the perspective of the assumptions he holds.

Central Issues Depend on the Theoretical Context

The issue concerning how much we ought to sacrifice for others in need becomes one of the central questions due to the theoretical context. Hooker's theory holds that we have the duty to help others in need while the theory also holds that we have special obligations toward particular people with whom we have some special relationships. Then, the question about how much we ought to give up our own goods for strangers in need is an obvious question we need to ask. What this shows is that the central questions Hooker's theory tries to answer are dependent on its theoretical tradition. Because of certain views the theory holds, the question concerning how much we ought to give up our own goods for helping others becomes an issue. The social intuitionist model might not explain why this particular issue becomes a central issue in Hooker's moral theorising since it is hard to see how mere emotional responses to individual cases lead us to considering this issue.

It also seems that Hooker discusses this issue due to the fifth criteria for assessing moral theory, namely that moral theory should provide an answer to the question about which

we are not confident enough. Hooker's choosing this issue may be partly caused by his referring to this criteria, not by his initial thought that the issue is important.

It is worth noting that this issue is a central question for all major normative theories. The fact that this issue is one of the central issues in normative ethics reflects that all normative theories do presuppose that we should help others in need while we have the right to use our property for our own goods. Because of these two theoretical assumptions, the issue becomes one of the central problems in ethics in general.

The Standard of Confirmation Depends on Background Assumptions

Hooker's explanation of wary rule-consequentialism clearly shows that the standard of confirmation of Hooker's theory is importantly dependent on existing theoretical assumptions. If there are two competing hypotheses both of them are likely to produce the expected value, the one which is closer to existing practices than the other is chosen. This is analogous to the theory-choice procedure of science: in science, a hypothesis whose theoretical assumptions are relatively similar to the ones of existing theories is chosen as a simpler theory even if there is a competing theory whose theoretical assumptions are different but capable of explaining relevant observable phenomena.

There is the Mutual Relation between Old Principles and Revised Principles

This theory-dependent feature is, again, seen in Hooker's a rule-consequentialist answer to the question concerning how much we ought to give up our own goods for others. Hooker's theory starts with the assumption that the internalisation of the moral code that we ought to help others in need generally produces good consequences. Hooker's rule consequentialist answer to the question gives a more detailed account of this assumption. The old principle used to say that we ought to help others in need while the revised principle gives a more specific account of this moral duty: we are required to help those in greater need, especially the worst-off, even if the personal sacrifices involved in helping others add up to a significant cost to us. This shows the mutual relationship between Hooker's old theory and revised theory. Again, it is hard to see how this sort of mutual relation can be seen if moral theorising is merely our making initial responses to given cases.

To sum up, my elucidation shows that Hooker's theory has certain theory-dependent features analogous to the theory-dependent features of science. What does this result show? This result shows that there is one real example in contemporary normative ethics which actually supports the second premise of the theory argument. In a sense, this is still a thin defence for the theory argument. It might be the case that other major normative ethicists do not exhibit any theory-dependent feature as Hooker does and in fact those other theories do not

support the second premise of the argument. As I indicated in the previous section, this is an empirical question and cannot be settled unless we do some real empirical work on the way normative ethicists build up their theories. In any case, it is still safe to say that in this section I have managed to present a case of a leading contemporary normative ethicist which supports the second premise of the theory argument. Thus, the defence provided here may be thin, but not insignificant.

7. Conclusion

In this paper, I considered whether recent empirical studies of moral judgements undermine the theory argument which is one possible way to defend naturalistic moral realism. I concluded that recent empirical studies do not undermine the theory argument. I gave three responses to the empirical objection against the theory argument. First, I argued that even if the recent studies support the social intuitionist model, this does not undermine the second premise of the argument. Second, I argued that the empirical findings shown are not supposed to undermine the second premise of the argument since these studies are about the psychology of lay people. I argued that these empirical studies need to be about the psychology of expert moral theorists. Third, I investigated whether the ways real normative ethicists build up their theories support the theory argument. I chose Brad Hooker's consequentialism, and claimed that we can see certain theory-dependent features analogous to the ones of science in Hooker's theorising. This

may be empirical evidence for the second premise of the argument, though this piece of evidence may be not conclusive.

The conclusion of this paper supports our initial thought that naturalistic moral realism goes hand in hand with the recent trend in philosophy. The more we get empirical data relevant to morality such as the ones concerning our moral judgements, the more we may be given reasons to believe naturalistic moral realism. For the defence of naturalistic moral realism, philosophers need to consider the proper implications we can draw from the empirical findings we have got, but this just shows that they can refer to certain empirical evidence for defending their position.

References Cited

- Boyd, R. 1988. How to Be a Moral Realist. in: Sayre-McCord, G. ed. *Essays on moral realism*. Ithaca, N.Y: Cornell University Press, pp. 181-228.
- Boyd, R. 2002. Scientific Realism. *Stanford Encyclopedia of Philosophy* [Online] 2009(03.10.2009). Available at: <http://plato.stanford.edu/entries/scientific-realism/> [Accessed: 12.01.2012].
- Brink, D. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Chaiken, S. et al. 1996. Beyond accuracy: Defense and impression motives in heuristic and systematic information processing. in Gollwitzer, P.M. and Bargh, J.A. eds. *The psychology of action: Linking cognition and motivation to behavior*. New York: Guilford Press, pp. 553-578.
- Chen, S. et al. 1996. Getting at the truth or getting along: Accuracy - versus impression -motivated heuristic and systematic processing. *Journal of*

- Personality and Social Psychology* 71, pp. 262-275.
- Chonabayashi, R. 2012. *A Defence of Metaphysical Ethical Naturalism*. PhD Thesis, Cardiff University [available at <http://orca.cf.ac.uk/26860/>] (accessed 2015/11/30).
- Collins, R. 2000. Reflexivity and Social Embeddedness in the History of Ethical Philosophies. in Kusch, M. ed. *The Sociology of Philosophical Knowledge* Berlin: Springer, pp. 155-178.
- Cooper, J. 2007. *Cognitive Dissonance: Fifty Years of a Classic Theory*. London: Sage Publications.
- Darley, J. M. and Berscheid, E. 1967. Increased Linking as a Result of Anticipation of Personal Contact. *Human Relations* 20, pp. 29-40.
- Darwall, S. et al. 1992. Toward Fin de siècle Ethics: Some Trends. *The Philosophical Review* 101(1), pp. 115-189.
- Doris, J and Plakias, A. 2008. How to Argue about Disagreement: Evaluative Diversity and Moral Realism. in Sinnott-Armstrong, W, ed. *Moral Psychology vol. 2*. Cambridge, Massachusetts: The MIT Press. pp. 303-332.
- Griffin, J. 1996. *Value Judgement: Improving Our Ethical Beliefs*. Oxford: Clarendon Press.
- Haidt, J. 2001. The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Modal Judgment. *Psychological Review* 108(4), pp. 814-834.
- Harman, G. 1986. Moral explanations of natural facts. *Southern Journal of Philosophy*, vol. 24, pp. 57-68.
- Harman, G. and Thomson, J. J. 1996. *Moral Relativism and Moral Objectivity*. Cambridge, Mass: Blackwell.
- Hooker, B. 2000. *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality*. Oxford: Oxford University Press
- Kohlberg, L. 1969. Stage and sequence: The cognitive - developmental approach to socialization. in: Goslin, D.A. ed. *Handbook of socialization theory and*

- research. Chicago: Rand McNally, pp. 347-480.
- Leiter, B. 2001. Moral Facts and Best Explanations. *Social Philosophy and Policy* 18, pp. 79-101.
- Lord, C. G. and Ross, L. 1979. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology* 37, pp. 2098-2109.
- Morgan, S. 2006. Naturalism and Normativity. *Philosophy and Phenomenological Research* 72(2), pp. 319-344.
- Mill, J. S. 1871. *Utilitarianism* Reprinted in Miller, D. E. ed. 2002. *The basic writings of John Stuart Mill: on liberty, the subjection of women, and utilitarianism*. New York: Modern Library.
- Miller, R. W. 1985. Ways of Moral Learning. *The Philosophical Review* 94(4), pp. 507-556.
- Rea, M. 2006. Naturalism and Moral Realism. in: Crisp, T. et al. eds. *Knowledge and Reality*. Dordrecht: Kluwer.
- Schwitzgebel, E. 2009. Do Ethicists Steal More Books? *Philosophical Psychology* 22(6), pp. 711-725.
- Schwitzgebel, E. and Rust, J. 2009. The Moral Behaviour of Ethicists: Peer Opinion. *Mind* 118, pp. 1043-1059.
- Shafer-Landau, R. 2003. *Moral Realism: A Defence*. Oxford: Clarendon Press.
- Sinclair, N. 2011. The Explanationist Argument for Moral Realism. *Canadian Journal of Philosophy* 41(1), pp. 1-24.
- Sturgeon, N. 1985. Moral Explanations. in Copp, D. and Zimmerman, D. eds. *Morality, Reason and Truth*. Totowa, NJ: Rowman & Littlefield, pp. 49-78. Reprinted in: Rachels, J. 1998. ed. *Ethical Theory 1*. Vol. 1. Oxford: Oxford University Press, pp. 180-210 - to which page references refer.
- Sturgeon, N. 2005. Ethical Naturalism. in: Copp, D. ed. *The Oxford Handbook of Ethical Theory*. Oxford: Oxford University Press, pp. 91-121.

Sturgeon, N. 2006. Moral Explanations Defended. in: Dreier ed. *Contemporary Debates in Moral Theory*. Oxford: Blackwell, pp. 241-262.

Turiel, E. 1983. *The Development of Social Knowledge: Morality and Convention*. Cambridge: Cambridge University Press.

ⁱ An exception is Doris and Plakias (2008) in which they explicate Boyd's argument as the theory argument. For an attempt to defend the theory argument, see Chonabayashi (2012).

ⁱⁱ I thank Nick Shackel for the discussion we had on the issues related to this section on which the argument I present here is built.

ⁱⁱⁱ At the moment, one available empirical study concerning the history of ethics indicates that there are certain theory-dependent features which played some essential roles in the revision procedure of ethical theories. See Collins (2000).

^{iv} Hooker is neutral about whether our moral judgements are beliefs or the expressions of our non-cognitive states.

和文要旨

自然主義的道徳的実在論と近年の社会心理学における知見について

蝶名林 亮

近年、経験科学で得られた知見を活用して伝統的に論じられてきた哲学的諸問題に関して新たな考察を示そうとする動向が多く見られ、その傾向性は高まってきている。「実験哲学 (experimental philosophy)」と呼ばれる近年の動きはその一例である。

一方で、悪さ、正しさ、勇敢さといった道徳的性質 (moral property) は経験的方法によって探求することができ、これら道徳的性質は他の自然科学で探求されている自然的性質と同じような仕方で存在しているとする「自然主義的道徳的実在論 (naturalistic moral realism)」と呼ばれる立場が現代メタ倫理学において見込みのある立場として提唱され、様々に議論されている。

一見すると、この立場は近年の哲学における動向と相性がよいように見える。というのも、自然主義的道徳的実在論が主張していることは、道徳に関する知識が観察や経験に基づいたア・ポステリオリなものであるということであり、そのような知識を得るためには関係する諸科学において得られる道徳に関する経験的知見が必要だということになるように思えるからだ。

ところが、自然主義的道徳的実在論の擁護のための論証に詳細な検討を加えてみると、経験科学において得られた道徳判断に関する知見がこの立場に対して必ずしも好意的でない可能性が浮上してくる。近年の道徳判断に関する経験科学における研究としては、Jonathan Haidt による研究が有名だが、Haidt によると道徳判断は伝統的に想定されていたような理性的なものではなく、判断を下す者の直感的な反応に拠っているものであり、理性は本来考えられていたような道徳的真理を発見するためのものではないとされる。こ

のような見解は、自然主義的・道徳的実在論の擁護を目指す者にとって、厄介な問題になり得る。というのも、この立場を擁護するための論証の前提の1つに、道徳における探求は科学における探求と類似するものであり、一種の理性的な営みであるとの考えがあるからだ。この考えは社会心理学で示されている知見と衝突するよう見える。

もし自然主義的・道徳的実在論と近年の社会心理学において得られた知見が衝突するということになると、後者は前者に対する経験的な反論になり得る。本稿ではこの問題について考察し、自然主義的・道徳的実在論の擁護を目指す論者がどのようにこの衝突を回避することができるのか、検討していく。