Applied Ethics

Risk, Justice and Liberty

Center for Applied Ethics and Philosophy

Hokkaido University

Sapporo, Japan

Published by Center for Applied Ethics and Philosophy
Hokkaido University
N10 W7, Kitaku
Sapporo 060-0810
Japan

Edited by Center for Applied Ethics and Philosophy

Printed in Japan

ISBN978-4-9904046-7-3

Contents

Introduction ii

Contributors iii

Cyberwar and Just War Theory----------1
Matthew BEARD

Two Theses of Moral Enhancement---------- 13
Takeshi SATO

Arguments in Professional Ethics Concerning Justice and Autonomy with Regard to the Ill, Injured and Disabled---------- 25
Paul JEWELL and Evdokia KALAITZIDIS

The Unbearable Lightness of Personal Identity: Messages from Bioethics---------- 39
Cheng-Chih TSAI

Rethinking Survivor Guilt: An Attempt at a Philosophical Interpretation---------- 52
Satoshi FUKUMA

The Conflicting Terms of Environmental Justice: An Analysis of the Discourse---- 66
Morgan Chih-Tung HUANG

Epicureanism about the Badness of Death and Experientialism about Goodness--- 85
Fumitake YOSHIZAWA

Liberty and Freedom: The Relationship of Enablement---------- 96
Michael YUDANIN

The Professional Morality of the Documentary Filmmaker----------109
Wu-Tso LIN

Introduction

This collection of essays is the final summation of the Seventh International Conference on Applied Ethics held at Hokkaido University on Octover 26-28, 2012. The conference was organised by the Center for Applied Ethics and Philosophy, Graduate School of Letters, Hokkaido University (Sapporo, Japan).

The purpose of this collection is to bring together the wide-ranging papers on various fields of applied ethics presented at the conference.

It is our hope that this collection will contribute to further developments in research on applied ethics and promote our Center's mission, which is 'to bridge the gap between theory and practice'.

July 2013

Center for Applied Ethics and Philosophy

Hokkaido University
Sapporo, Japan

Contributors

Matthew BEARD is a Research Associate with the Centre for Faith, Ethics and Society at The University of Notre Dame, Australia and Managing Editor for Solidarity: The Journal for Catholic Social Thought and Secular Ethics. He is a PhD candidate at Notre Dame, and has published in international journals including the Heythrop Journal and the Journal for Medical Ethics.

Takeshi SATO, PhD, completed his doctoral degree (a thesis on the moral philosophy of R.M. Hare) at Hokkaido University, and is Research Fellow of the Japan Society for the Promotion of Science. He specializes in metaethics and utilitarianism, and is interested in a study of human enhancement.

Dr. Paul JEWELL teaches ethics in Health Sciences. He is the author of *Disability Ethics* Common Ground 2010 and *Teaching Ethics* Hawker Brownlow 2011. Dr Evdokia KALAITZIDIS teaches ethics in the School of Nursing and is the author of *Client centered nursing ethics: A reconstruction of professional ethics for professional nursing* Lambert 2009. Both are at Flinders University in Australia.

Cheng-chih TSAI is an associate professor at Mackay Medical College, Taiwan. He received his PhD in mathematics from Edinburgh University, UK, and his current research interests include modal logic, philosophy of language, philosophy of time, and metaphysical/evolutionary basis of ethics.

Dr. Satoshi FUKUMA is Assistant Professor, College of Community and Human Services, Department of Community Development, Rikkyo University, Japan. His current research focuses on the association between the meaning of work and the guaranteed income policy. He is author of *Kantian Constructivism in Rawlsian Ethics* (in Japanese, 2007) and one of Japanese translators of John Rawls' *A Theory of Justice*, rev. ed. (2010). His forthcoming publication is "Rawls in Japan: A Brief Sketch of the Reception of John Rawls' Philosophy" (University of Hawaii Press, *Philosophy East and West*, 64:4).

Dr. Morgan Chih-Tung HUANG is an assistant professor of public administration at National Open University in Taiwan. He received his higher education at National Taiwan University, where he received his BA in law and minor in politics and MA in national development. He earned his doctorate in STS (Science, Technology and Society) at the University of Edinburgh. His research interests have mainly been in environmental justice, the politics of regulation, and STS. He is currently researching and writing a book on environmental justice and social theory.

Fumitake YOSHIZAWA is a PhD student at the Graduate School of Humanities and Social Sciences, Chiba University, Japan. His research interests include philosophy of death, metaphysics, and value theory.

Michael YUDANIN is a PhD student at the University of Georgia, USA. His research interests include philosophy of freedom, Kantian and non-Kantian ethics, and Wittgenstein's language games as a framework for analyzing philosophical problems.

Wu-Tso LIN was born on June 17, 1970, in Taichung, Taiwan. He majored in philosophy in his school life as well as he was an amateur photographer and documentary filmmaker. After completing his master and PhD education at Tung-Hai University in 1993-2004, he began to teach the class about general education and some class about film studies at Central Taiwan University of Science and Technology from 2004 till now.

Cyberwar and Just War Theory

Matthew BEARD

1. Introduction

I walk outside into a storm; the wind is blowing and rain is pouring down. I open my umbrella but find that I am still getting wet. At this point, I am confronted with two possibilities: (1) although the umbrella was designed to keep me dry in inclement weather, the makers did not envision such severe storms and in such conditions the umbrella is inadequate; or (2) given the severity of the storm, I need to use the umbrella in a specific way in order to stay dry (perhaps to face it directly into the wind, or to grip it higher up the handle). If (1) is true, then the designers of this umbrella need to reconsider its structural design and make changes to deal with the new challenge severe storms pose; if (2) is true, then I need to think about the best way to use my umbrella under circumstance I am not familiar with.

In various schools of ethical thought, technological developments have been the storm to the umbrella of moral philosophy. Some have advocated wholesale changes to the structure of moral systems in light of new technological development, whilst others have suggested that existing systems are flexible enough to accommodate technological advances.

In the area of military ethics the most recent storm has been in the form of "cyberwar," against which the long-standing tradition called Just War Theory (JWT) has tended toward the latter view. Most just war theorists in the area argue that JWT is a broad umbrella which has weathered many storms; used rightly, it will also be able to accommodate cyberwar. However, Randall Dipert challenged this conception in 2010, suggesting that the challenges of cyberwar require, at the very least, augmentation to the existing principles of JWT, if not an entirely new ethical system altogether.[1] This article argues against Dipert's view, suggesting that JWT's existing principles are sufficient to meet the new challenges of JWT.

2. Cyberwar, defined

In 1993, John Arquilla and David Ronfeldt (1993) prophetically declared that "cyberwar is coming!" To say that history has proved them right would be uncontroversial; indeed, in another article published earlier last year, Arquilla (2012) cites several examples of recent cyberwar activities including the Russian invasion of Georgia in 2008, the crippling cyber attacks on Estonia in 2007, and the role of the cyber in the revolutions in the Arab Spring. Arquilla's 2012 paper seems to endorse the definition of cyberwar provided by he and Ronfeldt in 1993.

1 Although Dipert is a lone voice, he is a loud one: his article is the most widely-read in history of the *Journal of Military Ethics*. As such, his ideas require some response.

> Cyberwar refers to conduction, and preparing to conduct, military operations according to information-related principles. It means disrupting if not destroying the information and communications systems, broadly defined to include even military culture, on which an adversary relies in order to "know" itself. (Arquilla & Ronfeldt 1993: 30)

This view of cyberwar leads to an anachronistic view of the 13th century Mongols as proponents of cyberwar, and the unusual suggestion that "[c]yberwar may actually be waged with low technology under some circumstances" (Aquilla & Ronfeldt 1993: 32). However, this view of cyberwar as "[e]fforts to strike at the enemy's communications and ensure the safety of one's own" (Arquilla & Ronfeldt 1993: 34) is not popular today; more commonly accepted is the definition of "cyberattacks" presented by Dipert (2010):

> Cyberattacks belong to a large genus of all kinds of attacks on information systems. Such attacks include traditional counterespionage and disinformation campaigns, old-fashioned destruction of telephone lines, jamming of radio signals, killing of carrier pigeons [...] and so on. We can restrict ourselves here just to attacks on modern digital information systems, that is, on computers and computer systems: intentional damage to software, hardware, and the operations of information systems. Even the restriction to digital is not especially crucial. (Dipert 2010: 386)

Also relevant is the Oxford English Dictionary definition (to which Dipert refers): "The use of computer technology to attack an organization, state, etc.; *esp.* infiltration or disruption of computer or other information technology systems for strategic or military purposes (OED, 2012). For my purposes here, I will adopt a view of cyberwar that reflects both Dipert and OED: cyberwar as *the use of computer software and technology by one nation to attack the governmental or civilian information systems of another nation*. Under Dipert's understanding, the bombing of (for instance) a building hosting a nation's social security infrastructure would count as cyberwarfare, but such actions seem to fall under the category of conventional warfare. *Cyber* warfare, as both the OED's and my definition notes, is war taking place in another realm; thus, the means by which cyberwar is conducted should be limited to digital — viz. software-based — methods.

However, the OED definition is also problematic because it extends cyberwar to include non-state attacks. However, even leaving aside the complex debate as to whether non-state actors are able to wage war, this definition would include actions better described as "cyber vandalism." For instance, the "hacktivist" group Anonymous' shutting down of US Department of Justice and FBI websites in protest of SOPA in January, 2012; or acts of "cyberterrorism" (Lewis 2002), of which there have been thankfully few instances. Such actions are clearly not acts of war, but they do fit within the OED umbrella. For these reasons I adopt my above definition of cyberwar throughout the discussion here.

Amongst (notably few) ethical commentaries of cyberwar, one of the few points

of consensus is that cyberwar poses a new challenge for those of us interested in the intersection of war and ethics (De George 2003; Lin et. al. 2004; Shackleford 2009; Rowe 2010; Dipert 2010). It is the position of some (Rowe 2009) that cyberattacks are incompatible with ethics, whilst others are more positive about the prospects of the moral use of cyberwar (De George 2003; Dipert 2010). However, whichever position one takes, it seems apparent that there are issues emerging from cyberwar which—at least on the surface—challenge traditional conceptions of JWT. These issues contribute to the view of cyberwar as a new type of war itself: "cyberwar signifies a transformation in the nature of war (Arquilla & Ronfeldt 1993: 31). However, we would do well—as others do (Rid 2012)—to treat this notion with a healthy dose of suspicion.

3. The claim that cyberwar is problematic for just war theory: A response to dipert

An emerging trend in military analysis (including military ethics) is the view that geographical and technological advancements are leading to changes in the nature of war (van Creveld 2004; Fleming 2004; Gray 2005; Gray 2010).[2] Naturally, the developments in cyberwar have led to similar inclinations. Under this view, cyberwar is a *bellum novum*, a new war, which "differ[s] from previous forms of warfare (Dipert 2010: 384).[3] This position in itself is interesting, but more interesting is the Dipert's knock-on view that cyberwar is a new type of war that traditional principles of JWT cannot adequately deal with, and that these principles therefore require augmentation. This section will focus on exposing and developing Dipert's argument, which is—as far as I'm aware—the only example from the philosophy corpus arguing for the inapplicability of JWT.

Dipert presents four major arguments to suggest that cyberwar is a new type of war, from which he goes on to suggest that it cannot be governed by traditional JWT —which Dipert understands to be defined by its doctrine of just cause (*casus belli*) (395)—unless JWT is modified in two ways:

> In a *moral* dimension toward a general notion of harm, especially to national vital interests [...] and in an *ontological* dimension, with a focus away from strictly injury to human beings and physical objects toward a notion of the (mal-)*functioning* of information systems, and the other systems (economic, communication, industrial, production) that depend on them. (386)

The four arguments Dipert presents for cyberwar's status as *novum bellum* inform these two arguments, and thus refuting them provides a means of responding Dipert's changes. The four arguments are as follows: (1) Cyberwar does not

2 Brahms composes a much longer list of literature arguing for the changing nature of war at n.1 of his article.

3 All references in this section, unless otherwise noted, are from Dipert (2010).

require violence, lethal force or even permanent damage to objects (385); (2) The harms produced by cyberwar are usually of a non-physical nature, or where there is physical harm, it is indirect and may occur some time after the actual attack is performed (395); (3) Cyberattacks do not fit the description of aggression described in contemporary JWT (395); and (4) identifying the perpetrator of a cyberattack is extremely difficult, which he calls "the Attribution Problem" (393).

I will begin by discussing (1), but in so doing, will also be developing the argument for (3), which is informed by (1). Dipert notes that JWT has been motivated by concern about "the lethality and massive destructiveness of war." This makes cyberwar a curious case because "cyberwar often will not be like that" (386). In fact, as Dipert's taxonomy of cyberharms shows, cyberattacks may be neither lethal, nor broadly destructive. This useful taxonomy relies on a distinction in computer science between different entities: data, algorithms, the algorithm's application to the data, and the hardware (398). Only the last of these—the damage of hardware—counts as damage which might be called 'widespread destructiveness'; harm to digital systems or data is not incorporated in existing international law or principles of Just War.[4]

Dipert examines the United Nations (UN) Charter's definition of aggression as "the threat or use of force against the territorial integrity or political independence of any state." (UN Charter 1945: Ch.1 Art. 2 Sec. 4) and Walzer's (2006) similar definition as "[e]very violation of the territorial integrity or political sovereignty of an independent state" (Walzer 2006: 52), observing that both of these seem to place "armed" attack at the heart of their accounts. Obviously cyberwar's status as armed attack is controversial, and it is "not obvious" why we *should* consider them to be armed attacks (396) as this seems "literally understood, to designate soldiers using 'arms', roughly, as artifacts for inflicting injury, death, or causing physical destruction of objects" (395). Existing theories of *casus belli* limit the type of attack worthy of the response of war to physical attacks. How then should we respond to cyberattacks? Can we ever respond to a cyberattack with a conventional strike? These are questions that JWT struggles to answer.

The next—related—concern is that cyberharms are not physical in nature. How can JWT deal with this when, for instance, "[a] cyberattack does not involve intrusions into the territory or airspace by soldiers or even by physical objects?" (397) Such an incursion does not look like the paradigmatic form of aggression (for Dipert, Pearl Harbour). The Russian cyberattack on Estonia prompted the Estonian

4 There is, however, a sense in which this type of damage *is* physical in nature: although software changes do not correlate to the *destruction* of any physical object, they do cause a change in the storage systems of that object, where the data physically manifests in magnetic signals on the hard drive. Thus, the deletion or alternation of data *does* correlate to a physical *change*. One can destroy data from within a computer, but also by exposing the hard drive to substantial magnetic energy. I do not think this undermines Dipert's argument, but it is worth noting that there is a physical component to the harm here. Cyberattacks are nuanced type of attack, but they are not entirely non-physical. The equivalent would be perhaps to say that psychological harms are not non-physical insofar as they correspond to changes in parts of the victim's brain/mind.

Prime Minister to ask "What's the difference between a blockade of harbors or airports of sovereign states and the blockade of government institutions and newspaper websites?" (Rid, 2012). JWT's response, Dipert seems to think, would be the presence of military forces.

Dipert's contention is that JWT's interest in physical attack and aggression makes its application to cyberwar unclear (397). He asserts that most just war theorists tend to follow a "view of aggression or attack as invasion or destruction" (396), citing Nicholas Fotion as evidence of this. However, this is to misrepresent Fotion, who actually posits a new (and controversial) view of *casus belli* called the "multiple reasons process" (Fotion 2007: 76) which allows "wars triggered by a series of small acts of aggression, assassinations, sabotage, systematic harm done to people[...]" (Fotion 2007: 77). Fotion, who Dipert takes to be representative, actually holds that there *are* just causes for war beyond aggression.

However, it might be argued that few just war theorists subscribe to Fotion's multiple reasons process", and therefore Dipert's description of JWT still holds, even if his description of Fotion does not. This objection fails to hold muster though, as more traditional theorists–including Michael Walzer and Brian Orend (2006), both of whom Dipert cites — are more interested in aggression *vis a vis* a breach of *states' rights* (or, as classical theorists like di Vitoria would put it, "a wrong received" (di Vitoria 1991: 319) than in aggression *vis a vis* physical attack. This is the key fact that Dipert ignores: armed attack is not the key condition of aggression, it is the breach of territorial integrity and/or political sovereignty that constitutes aggression in international law and Walzer's JWT. The broad attack on the suitability of JWT based on the non-physical nature of cyberharm is insubstantial. The proportionality criterion of *jus ad bellum* does not stipulate that the *type* of harm suffered be reflected in retaliatory action, but that the *degree* of harm be reflected. What is significant is that (1) harm is suffered, (2) that there is enough of it, and (3) that it is harm against either the citizens of a state or the state as a whole.

These two ideas: proportionality and (more importantly) defining aggression relative to states' rights give a means of defending JWT against Dipert's grievance. In responding to the Estonian Prime Minister, JWT might say the following:

> There may be no difference, in principle, between these two types of blockades. Indeed, both breach rights to which you are entitled as a state. However the duration of the 'blockade' was roughly a total of 3 hours and 30 minutes; a conventional blockade that lasted less than a day would no more warrant military action than would this 'cyberblockade'.

It is also worth noting that Dipert rightly shows that the United Nations Charter limits aggression to armed attack (395), I would suggest that this points to the inadequacy of existing international law to deal with the new challenge of cyberwar, not the inadequacy of JWT. Furthermore, I would suggest that the international law's adoption of a broader definition of aggression (such as JWT's) would help resolve such problems. Here it seems worth noting that the moral augmentations that Dipert calls for are unnecessary because JWT already has a "general notion of harm"

centred around states' rights.

I will return to the question of how one can morally respond to cyberattacks.

Moving briefly to "the Attribution Problem", this is a nonissue for JWT. The Attribution Problem, Dipert suggests, is based in the deeper epistemic question of "how much justification or evidence is necessary concerning the threshold conditions for morally going to war?" (393) This question, he alleges, "[has] been ignored by most theorists of the morality of war" (393). However, whether or not this is true (and I suspect it might be true of contemporary commentators, but not of more classical sources) does not mean it is a question outside the domain of JWT. For instance, Grotius argued that in cases of uncertainty about the justice of one's cause, one should err on the side of peace (Grotius 2006: Bk. II, XXII.V), meaning, in short, that one ought not to go to war.

This epistemic concern is also not unique to cyberwar: indeed, the "fog of war"—the inherent uncertainty in dealing with military issues—is acknowledged by a number of theorists (Walzer 2006: 281; Orend 2006: 44, 114; Elshtain 2003: 102). Another issue in which it is particularly prevalent is in discussions of the morality of preventive war; a point which Dipert acknowledges (393). In this area debate continues, but the international law seems to require that an impending attack be *certain* before legitimizing the attack (Bothe 2003: 232; Zedalis 2005: 214), and certainly classical just war theorists maintained that one must be certain of a forthcoming attack before military action is justified (di Vitoria 1991: 316; Grotius 2006: Bk. II, I.I). Even Walzer (2006), who permits some preventive actions ("anticipations"), requires an enemy show "manifest intent to injure" (Walzer 2006: 81), indicating that without at least *near*-certainty, one's cause will not be just.

The point here is simply that JWT is not ill-equipped or inadequate in dealing with the Attribution Problem *from a theoretical standpoint*. It will, no doubt, be very difficult for states wanting to act well to respond to secretive cyber attacks, just as it is difficult for those same states to justly respond to terrorism; but to say that because of this JWT cannot fit the Attribution Problem under its theoretical umbrella is misguided. Certainly, the relative-anonymity of cyber attacks is problematic, and solutions should be suggested (including Dipert's own) (393; Dipert 2006), but the important point is that these solutions can come from *within* the logic of existing JWT.

Returning now to the question of whether cyberattacks can be responded to with conventional (or cyber) attacks, two things might be said: first, once again, JWT can answer this question within already existing frameworks, and secondly, that Dipert's own response to the question comes very close to doing so. Dipert lists four conditions for justified conventional or cyber attack in retaliation to an initial cyberattack.

(1) The attack of C on B was unjust and substantial.
(2) The source of this attack by C was, with overwhelming likelihood, ordered or permitted at the highest levels of a government.
(3) Reasonable measures had been taken by nation B to defeat or minimize the cyberharm that a hostile nation or other non-state cyberattacker (black or

grey hat hackers) might cause.

(4) The expected damage to the enemy (C) is likely to be commensurate to the damage B has suffered, or is the minimum necessary to stop continuing cyberattacks.

Conditions (1) and (4) are mirrored exactly by existing criteria in JWT; specifically, *casus belli* and proportionality. Condition (2) does not mirror the JWT of classicists such as Grotius, who would insist on absolute certainty before war could be entered, but the requirement that engaging in war is treated with at least a great deal of caution is consistent with the general spirit of JWT. (And nevertheless, JWT could simply rewrite (2) as "the source of this attack by C was, certainly, ordered or permitted at the highest levels of a government" without a great deal of difficulty). So, the only really novel contribution Dipert appears to make to the ethics of cyberwar is condition (3), what could be called 'The Principle of Cyber Self-Protection'.

This principle is based in the view that a responsible state ought to be protected against the many cyberthreats present in the modern world. "Some condition like this [...] seems required in the case of cyberwarfare because civilians are in possession of tools that are as destructive as those nations can organize [...] and because one can expect occasional attacks or mischief from them" (401). However, it takes only a little thought to see that this principle is a bad one. The state's failure to fulfil one obligation does not render it immune from all its rights, including the right to wage war against an unjust assailant. It might be established that prudence demands that the modern state be equipped with cyber defence, but the logical jump to say that failure to do so means a state may *not* retaliate from unjust attacks is counter-intuitive.

Imagine a similar principle, 'The Principle of Protection from Terrorism.' One might argue that the rise of extremism has meant that for the last few decades, states have occasionally been threatened with terrorist attacks. States should, therefore, be equipped to overcome such attacks in the present day. However, the suggestion that the United States had no just cause for retaliating to the 9/11 attacks because of a failure to adequately protect herself is ludicrous, and contrary to the JWT belief—based in Vitoria (1991: 319)—that states have the right to retaliate after being attacked, even if it was ill-prepared for, is still grounds to punish that evil, and failure to take all reasonable means of protection does not render the right of retaliation or punishment moot. Dipert uses a legal analogy—one of suing a person who intentionally, but to my knowledge, polluted my drinking water—to justify his principle, "I might indeed win a civil suit (or a criminal complaint) but the amount of the settlement would almost certainly be reduced by my failure to take reasonable precautions" (402). But the analogy is false: restitution for damages and punishment for crime are not analogous; a better analogy would be if Dipert was forbidden from suing the intentional polluter at all. It seems that the only complete deviation Dipert makes from JWT results in a bad principle, which provides us with good grounds for believing that JWT is, at the very least, better than any existing alternatives at dealing with cyberwar.

4. Just cyberwar theory

Dipert's approach suggests that JWT's 'umbrella' must *change* in order to accommodate cyberwar, but in the last section I suggested that Dipert is heavily reliant on the same theory he seeks to jettison, and for this (and other) reasons, I believe his approach fails. In this last section I want to show how a more simply application of some principles of JWT might approach some of the contentious issues cyberwar provokes. I should note that this does not aim to be a systematic treatment of JWT and cyberwar; this would be a much larger project. Rather, I aim to show how JWT might be used to approach such issues, as well as attempting to contribute to the growing debate around these areas by adding a new voice. This section will discuss how cyberwar-specific issues pose resolvable challenges to JWT, following the traditional division of the subject into *jus ad bellum*, *jus in bello* and *jus post bellum*.

4.1 *Jus ad bellum*

It will not seem immediately clear to all that cyberwar poses any significant challenge to the doctrine of *jus ad bellum*; for those who believe cyberwar to refer only to the means by which a war is fought, cyberwar will be limited to *jus in bello*. However, such an approach is, I believe, mistaken. Serious questions must be asked about the relationship between cyberwar and *jus ad bellum* concepts; specifically, just cause and proportionality.

Perhaps the most interesting challenge is posed to just cause: as Dipert notes, cyberwar attacks do not reflect traditional acts of aggression which modern JWT sees as synonymous with just cause (the logic being that if a nation has committed aggression, then the responding cause is just). Does a state-sanctioned and organised attack on the financial market of another nation count as aggression, and therefore as a just cause? If so, is it justifiable to respond with conventional weaponry?

To the question of whether a state-sanctioned cyberattack counts as aggression, I believe the answer is certainly yes. There seems to be no relevant difference between disabling financial markets via the internet (or some other digital medium) and disabling them by employing a spy to manually cut cables, the intention is the same: malicious damage to a nation's infrastructure. To Dipert's claim that the harm that has been done is not *physical*, the obvious response is to point to the long term consequences of such an attack: confusion, panic, fear, people's investments lost, diminished faith in markets, unemployment, etc., and point to the obvious 'physical' harms these generate. It is likely that less blood will be shed, but the overall extent quality of harm may in fact be much greater than in some minor conventional wars.

In light of this, does being a victim of a cyberattack constitute a just cause for war? Perhaps; it certainly warrants some response. The first, and most obvious, would be the repelling of the assailant and the restoration of affairs to the way they were prior to the attack (assuming this state of affairs was just to begin with). A nation is certainly justified in doing this digitally (say, through a counter DDoS attack), just as it would be justified in using force to capture and detain a team of

spies sent to fulfil the same purpose. Further, a nation would be justified in "declaring war" on their assailant (if they could be confident of the origin of the attack), where declaring war means ceasing trade relationships, closing embassies, demanding compensation and publicly condemning the attack. But would it also be permitted in actually employing force (conventional or cyber) against the opposing nation for the attack?

In short, no. The reason here is the same as it would be if there were a physical attack. The proportionality requirements of JWT require that war be fought as minimally as possible: if my nation is invaded, I am justified in using force to repel the attackers and even taking means to ensure they will not attack again, such as installing border defenses, and even crippling their military to delay the possibility of another attack, but I am not justified in bombing my enemies capital city. My just cause ends once my lands are safe and my security assured (as best it can be). The same is true for a cyberattack: I could not respond by counter-attacking the financial systems of the aggressor nation, because this exceeds the end of the war, viz. the restoration of the *status quo ante bellum*, as Walzer explains:

> [M]any war aims can be achieved well short of destruction and overthrow. We need to seek the legitimate ends of war, the goals that can rightly be aimed at. These will also be the limits of a just war. Once they are won, or once they are within political reach, the fighting should stop. (1977: 110)

It seems to me that the aims of a war in response to a cyberattack, particularly in the case where the cyberattack is a "one off" are limited to removing the threat and correcting the harms that have been inflicted. Neither of these goals seem to justify a counter-attack, and certainly do not justify a conventional attack (which Dipert allows for). If an excessive cyber-response or conventional attack were pursued, the cause would no longer be justified. Given this, military responses to cyberattacks are largely limited to defensive technologies: ways of preventing attacks and quickly ending them if they occur. This seems to validate De George's suggestion "that a nation has the obligation to develop what defenses it can to protect its people" (De George 2003: 185). For instance, "if one's potential enemy is working on encryption, firewalls and other secure means of communication, then legitimate self-defense interests require that one do at least as much, and preferably more." (185) This claim seems to be the basis of Dipert's view that if a nation fails to protect itself digitally, then it forfeits the right to a retaliatory cyberattack. However, De George usefully highlights that the failure to develop cyber defences is a failure to a nation's own citizens, who it leaves at risk of harm; thus it should afford *those citizens* special rights (to compensation, for example), rather than affording protection from retaliation to aggressive cyber attackers.

Of course, this is all in the case where attacks were initiated by cyber methods. In the case where one chooses to respond to a conventional strike with cyber-based technologies, I believe such a response would be justified almost always, as long as it adhered to the *jus in bello* requirements which below. In fact, as long as such an attack met the *jus ad bellum* 'probability of success' requirement, this response

would be particularly meritorious insofar as it reduces the amount of bloodshed likely to occur.

4.2 *Jus in bello*

The two principles that most JWTs tend to agree govern *jus in bello* are discrimination, which governs who is a legitimate target of attack, and proportion, insisting against the use of "excessive force". Perhaps the most interest in cyberwar and *jus in bello* has been concerning discrimination; specifically, civilian immunity.

In 2010, a virus worm known as 'Stuxnet' was discovered. It targeted Iranian nuclear facilities and was believed to have been designed by Israeli and US operatives. The virus was disseminated widely, infiltrating a vast amount of computers (it is possible that the PC I am writing on now is infected with the virus), but would not cause any damage unless a small, industrial-use Siemans product (used in Iranian facilities) is present (Schneir 2010). If we assume my PC has been infected with Stuxnet, then I have been the innocent victim of a cyberattack; however, I am clearly not the victim the same way a citizen living through the firebombings of Tokyo was a victim. Unlike that poor soul, I have not been harmed by my attack. I do not believe such an attack breaches *jus in bello* protections of civilians: although I have been "attacked", I have not been harmed.

Contrast this type of civilian targeting to the way civilians are targeted in an attack that disables the electricity to a city. If we assume the city possesses some military facilities which are the primary target, then the moral logic works in a similar manner to that which justifies the Stuxnet attack, but this attack is not justifiable, for reasons identified by De George.

> Consider a city like New York. To deprive it of electricity would be to paralyse it. And if all the circuits were burned out and had to be replaced, the task would be enormous. Add to that the destruction of the communications systems, the transportation system, and all the private and business computers. The city would stop functioning except on the most primitive level, and hence the effect on innocent civilians would be devastating. (2003:188)

This type of attack deals widespread harm to civilians, and thus violates the requirements of discrimination. In a sense, the Stuxnet attack is by far the more direct, and personal, but the immorality of the second is immediately clearer in terms of the severity of the harm dealt. (Although this is not to say that harm is the *only* measure of civilian immunity; if the Stuxnet virus was collecting data and spying on every PC it infected, it too would be morally condemnable for its breaches of privacy and treatment civilians as instrumental to the achievement of military ends).

4.3 *Jus post bellum*

On the matter of the resolution of war, there is very little to say specifically about cyberwar, except that any technology (such as viruses) installed on hostile technology must be removed, rather than left for posterity. To leave Trojan horses

or other digital surveillance technology after a war's completion is to anticipate further war in the future, and undermines efforts at peace: as Kant says in the first Preliminary Article of *Perpetual Peace*: "No treaty of peace that tacitly reserves issues for a future war shall be held valid." (Kant 1983: 107).

5. Conclusions

Cyberwar is one of many developments in warfare that have been claimed to have changed the nature of war. Although Dipert argues that the new developments of cyberwar are beyond the reach of traditional JWT, he ignores the breadth of those principles. Furthermore, the few unique points he makes prove to be inadequate to govern the ethics of cyber warfare, casting further doubt on his assertion.

By contrast, it seems that the principles of JWT applied thoughtfully to cyberwar *do* hold, and are demonstrative of the applicability of JWT to particular (and as yet unforseen) developments in warfare. Cyberwar is a *strategic* development in warfare, not a change to its essence. The principles laid down by JWT reflect the essence of war, and therefore will only be challenged to a change in that essence, not by developments to superficial aspects of its practice. JWT is an umbrella which holds up against most storms, if it is used effectively.

References

Anscombe, G.E.M. (1965). War and Murder. In Wasserstrom, R. (Ed.) (1970). War and Morality. California: Wadsworth.

Aquinas, T., Fathers of the English Dominican Province (1920). The Summa Theologica of St. Thomas Aquinas. http://www.newadvent.org/summa/

Arquilla, J. (2012). Cyberwar is Already Upon Us. Foreign Policy. 192.

Arquilla, J., Ronfeldt, D. (1993). Cyberwar is Coming!. Comparative Strategy. 12(2), 141-165.

http://www.foreignpolicy.com/articles/2012/02/27/cyberwar_is_already_upon_us

Augustine, Bettenson, H. (2003). The City of God. Victoria, Australia: Penguin Books.

Bethke Elshtain, J. (2003). Just War Against Terror: The Burden of American Power In A Violent World. New York: Basic Books.

Bothe, M. Terrorism and the Legality of Pre-Emptive Force. European Journal of International Law. 14, 27-240.

De George, R. (2003). Post-September 11: Computers, ethics and war. Ethics and Information Technology. 5, 183-190.

di Vitoria, F., Pagden, A. & Lawrence, L. (1991). Vitoria: Political Writings. Cambridge, UK: Cambridge University Press.

Dipert, R. (2010). The Ethics of Cyberwar. Journal of Military Ethics. 9(4), 384-410.

Fleming, B. (2004). Can Reading Clausewitz Save Us from Future Mistakes?. Parameters. Spring 2004. 62-76.

Fotion, N. (2007). War and Ethics: A New Just War Theory. London: Continuum.

Gray, C. (2005). How Has War Changed Since the End of the Cold War?. Parameters. Spring 2005. 14-26.

—— (2010). War: Continuity in Change, and Change in Continuity. Parameters. Summer 2010. 5-13.

Grotius, H., Reichberg, G., Syse, H. & Begby, E. (2006). The Ethics of War: Classic and Contemporary Readings. Victoria, Australia: Blackwell.

Kant, I. trans. (1983) Humphrey, T. *Perpetual Peace and Other Essays on Politics, History and Morals*, Indianapolis: Hackett Publishing Company.

Lin, P., Allhoff, F., Rowe, N. (2004). War 2.0: Cyberweapons and Ethics. Viewpoints. 55(3), 24-27.

Miller, R. (2009). Just War, Civic Virtue, And Democratic Social Criticism: Augustinian Reflections. The Journal of Religion. 89(4), 639-646.

Orend, B. (2006). The Morality of War. Toronto: Broadview.

Primoratz, I. (2007). A Philosopher Looks at Contemporary Terrorism. Cardozo Law Review. 29, 33-51.

—— (2011). Terrorism. Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/terrorism/

Rid, T. (2012). Cyberwar Will Not Take Place. Journal of Strategic Studies. 35(1), 5-32.

Ronzoni, M. (2012). Two conceptions of state sovereignty and their implications for global institutional design. Critical Review of International Social and Political Philosophy. 15(5), 573-591.

Rowe, N. (2010). The Ethics of Cyberweapons in Warfare. International Journal of Cyberethics. 1(1), 20-31.

Schneir, B. (2010). The Story Behind the Stuxnet Virus. Forbes. 10/07/10. Accessed: http://www.forbes.com/2010/10/06/iran-nuclear-computer-technology-security-stuxnet-worm.html 30/01/13.

Shackleford, S. (2009). From Nuclear War to Net War: Analogizing.

Cyber Attacks in International Law. Berkley Journal of International Law. 25(3), 193-251.

UN Charter (1945). The Charter of the United Nations.

www.un.org/en/documents/charter/chapter7.shtml.

van Creveld, M. (2004). Modern Conventional War: An Overview. Paper presented at the National Intelligence Council Conference, Washington, D.C. http://www.offnews.info/downloads/2020modern_warfare.pdf

Walzer, M. (2006). Just and Unjust Wars (4th ed.) New York: Basic Books.

Zedalis, R. Circumstances Justifying Pre-emptive Self-Defense: Thoughts Prompted by the Military Action Against Iraq. Nordic Journal of International Law. 74, 209-203.

Two Theses of Moral Enhancement

Takeshi SATO

1. Introduction

Today's biomedical technology can provide several kinds of human enhancement; enhancements of muscle, appearance, cognitive ability and character. Some argue that they are morally permissible, and some oppose them outright. Among enhancements, the most disputable technology of enhancement might be "moral enhancement" which aims at improving our morality. We could easily and directly be more moral beings by biomedical intervention.

The methods of moral enhancements that are currently dealt with are mainly intervention in our morally deleterious feelings, character traits or behaviors, as evinced in discrimination against and verbal or physical belligerence toward others. Despite disputes over definitions or the contents of morality, most people concede that such feelings are morally bad.

Some proponents of moral enhancement hold that we *must* strive to enhance our morality (strong moral enhancement thesis), and others hold that we *are welcome to try to* enhance our morality (weak moral enhancement thesis). According to the former thesis, our naturally given morality is ineffective in a world rife with morally tragic situations; therefore, we must modify our natural morality, especially feelings of empathy and fairness (Persson & Savulescu 2008). On the other hand, the weak moral enhancement supporters posit that it is wrong to regard all moral enhancements as morally wrong. They think that there are some morally permissible moral enhancements if certain conditions are met (Douglas 2008).

I think, however, neither a strong nor a weak argument for moral enhancement is sufficiently plausible. Therefore, the first section of this paper aims to discuss the feasibility of moral enhancement, while examining some arguments in its favor. In section 2, we will outline the current situation concerning moral enhancement. In section 3, we will consider two arguments for moral enhancement. Then, in section 4, we will examine whether the current forms of moral enhancement should be morally approved or not.

2. What is moral enhancement?

2.1 Aims of moral enhancement

Suffice it to say, the primary aim of moral enhancement is to enhance people's morality. But when we start to think about it from many angles, this aim seems to be difficult to attain. The first problem is *what do we mean by morality* and *what does it mean to be moral*. There are so many arguments, and some philosophers hold that morality solely relates to an individual's or a group's good will. Through complying

with the good will and categorical imperative, we can be moral agents. Some philosophers hold that morality relates to promoting the greatest happiness or desire fulfillment of the greatest number. Others contend that morality relates to a virtuous character; what kind of person one is. It is a very difficult problem of ethics to find any overlapping consensus among them.[1]

However, in the current context of moral enhancement, *moral* is often viewed optimistically to mean one who is *sympathetic, friendly, kind, just and fair -minded*. T. Douglas, who is a proponent of moral enhancement, contends that, "there are some emotions [...] whose attenuation would sometimes count as a moral enhancement regardless of which plausible moral and psychological theories one accepted"(Douglas [2011] p.470). Then, he considers that such emotions might manifest as a strong aversion to members of certain racial groups and violent aggression. So let's assume that morality relates to having some positive emotions like sympathy and friendliness and a lack of certain negative emotions like verbal and physical aggression, (I shall criticize this conception of morality in a later section).

The second issue is a so-called "*why be moral" problem*. Previous arguments between proponents and opponents of moral enhancement have typically focused on whether moral enhancement is morally acceptable or not. However, should we be moral? If we do not have any persuasive reasons to be moral, then we do not have any reason to enhance our morality. Furthermore, even if there could be a reason to be moral, what is it? If each theory supposes different reasons to be moral, in the end, they might not be talking about the same thing when they argue why moral enhancement is morally acceptable.

Though there are many arguments about how to deal with this problem, there seems to be two alleged reasons to be moral in the context of moral enhancement. One is to survive. We human beings are facing devastating world-scale problems, e.g. environmental problems, climate changes, poverty, and the threat of nuclear war. If we cannot deal with such serious global problems, our future seems to be bleak. To overcome our bleak future, we need to become somehow more moral beings, especially fair and just-minded (I shall call these reasons global reasons to be moral). The second reason to be moral is more personal; it is to live a good life, particularly by creating many good personal relationships. Our well-formed moral life essentially is based on good relationships with our family, friends, and neighbors. If we learn to be more sympathetic, friendly and kind, we should be able to easily create good relationships with the people close to us. Then, if we want to maintain such relationships, we have a reason to be more moral (I shall call this reason a private reason to be moral).

To summarize this section, the objective of moral enhancement is to enhance one's morality. The words "to enhance morality" mean to further some kind of feelings like sympathy and fairness. There are two alleged reasons to be moral, global reasons and private reasons.

1 cf. Scaefer [2011].

2.2 Means of moral enhancement

How can we enhance our morality? Traditionally, we do so through education. Philosophers like Socrates at *Agora*, Plato at *Akademeia* have provided education for young people and every teacher since ancient times has told his young students that they should be kind and fair to their friends. But future teachers might not have to say a word, if they have the luxury of dispensing a pill to enhance their students' morality. A modern way to make people moral is supported by application of biomedical technology, e.g. drugs, gene manipulation and brain machine interface.

For example, it is often said that oxytocin mediates our feelings of trust, sympathy and generosity (Douglas [2011], Persson & Savulescu [2012]). For mutual trust, it is essential to construct a good relationship with others, and this provides personal incentive for being moral. Oxytocin neurohypophyseal neuropeptide synthesized in the brain and released in the posterior pituitary, and in the central nervous system (Kovacs G.L., Sarnyai Z., Szabo G., [1998] p.945). It is naturally produced, so the risk of side effects is relatively low. SSRI (selective serotonin reuptake inhibitor) is also used as a moral enhancer (ibid.). It is said that it makes subjects more fair-minded and willing to cooperate with others. So receiving SSRI might serve global reasons to be moral which demand fairness.[2]

3. Arguments for moral enhancement

In this section, I will consider two arguments for moral enhancement. The first is a radical one that supports the thesis that we *ought to* enhance our morality. The second is more moderate and contends that we *may* enhance our morality. To reiterate, the former thesis is a strong moral enhancement thesis (SMT) and latter one is a weak moral enhancement thesis (WMT). A discussion of each theory's merits follows.

3.1 Strong moral enhancement thesis

SMT posits that there is a moral obligation to enhance our morality. It entails that someone who does not enhance or whose morality is not enhanced is blameworthy. This obligation and blameworthiness arise from a global reason to be moral. In fact, steady growth of our efficient technology confronts us with the possibility of global-size destruction, in the form of environmental disaster, and what is needed to avoid such a future catastrophe is not only certain scientific technology, but also cooperation from each of us. For example, we all have to reduce our use of fossil fuels, endeavor to recycle materials more and refrain from the misuse of advanced technology to treat such global problems.

Our current morality is, however, biologically restricted to those near and dear, and the near future. We often act from partiality, for example, preferring our children

2 Whether there are differences, especially morally relevant differences, between a traditional and modern way to enhance our morality is an important point of contention. Proponents of moral enhancement see the moral enhancer as just another instrument like a pair of glasses or a textbook, so they think it does not have any internal goodness or badness.

to others. And we also act on bias and deem the near future as more important than the distant future, and pursue smaller, immediate satisfactions rather than delay gratification to enjoy greater gains and pleasure in the long-term. These biases are a result of an evolutionary adaptation to the environment in which our ancestors had grown up in, and they used to fit well with their small society. However, "there is a widening gap between what we are practically able to do, thanks to modern technology, and what we are morally capable of doing" (Persson & Savulescu [2012] pp.106-107). Our naturally given capacity for morality can no longer fully meet the needs of our situation. Persson and Savulescu say that, "the moral shortcomings of humankind make the risks of catastrophic misuses of these [scientific] powers too great" (ibid. p.134). Consequently, we *must* overcome such a natural restriction of morality by using scientific technology (Persson & Savulescu [2011], [2012]). Some sorts of biomedical technology are considered to enable us to act more justly and fair-minded. According to Persson and Savulescu, "our point is just that the predicament of humankind is so serious that all possible ways out of it should be explored. Therefore, it is important that moral bioenhancement is not written off without good reason" (ibid. p.123).

Before proceeding to WMT, it should be noted that if we acknowledge SMT, we must make *people* more sympathetic and fair-minded in order to deal with catastrophe. Consequently, SMT has two important characteristics; (a) its targets are people who do not act in a sufficiently moral way and (b) moral enhancement must be done on behalf of *others,* since the effect of moral enhancement sometimes presents in future generations. In section 4, I will argue that these characteristics of SMT present some difficulties.

3.2 Weak moral enhancement thesis

In contrast to SMT, WMT just states that we *may* enhance our own morality, if we so desire (Douglas [2011]). So this does not imply that someone who has not enhanced his own morality is blameworthy. Since WMT contends that there are at least one or more situations that permit moral enhancement, there is no obligation to deliberately enhance oneself. This obligation arises often from private reasons to be moral. Let's suppose there is a man called Smith who is not sufficiently friendly with people close to him. In Smith's case, it is permissible to enhance his morality under the following conditions.[3]

1. Through moral enhancement, Smith will have better motives than he would otherwise have had.
2. Otherwise, he can expect to have bad motives.
3. The biomedical intervention will work by attenuating some of Smith's emotions.
4. The only effects of Smith's intervention will be to alter his psychology in ways necessary to bring about that he expectantly has better motives, and has no unexpected effects.

3 This formulation is originally taken from Douglas [2011] p.473. The words are slightly modified.

5. He can freely choose whether or not to enhance his morality, and if he chooses to do so, his choice will be based on the best possible reasons.

Hence, when Smith's aggression or lack of sympathy makes it difficult to construct and maintain good human relationships with his children, friends and neighbors, or to overcome an aversion to a certain racial group, moral enhancement should be morally permitted. In the case of the WMT, in contrast to SMT, its targets are not others, but the individual self, and it emphasizes rational self-determination to enhance one's own morality.

4. What is the problem with moral enhancement?

So far we have seen two main arguments for moral enhancement. In this section I shall examine both arguments and try to show that they are not provided with sufficient moral justification.

4.1 Strong moral enhancement thesis and global reasons

At first glance, SMT is apparently plausible. In fact, if we have no choice other than moral enhancement to survive future catastrophe, it seems to be true that we should enhance our morality (unless everyone altogether prefers the destruction of the whole world to the scratching of their finger). However, moral enhancement should still be considered as a last resort. Since there are at least four concerns about enhancing morality in such a situation, we should not easily invoke moral enhancement as a means to avoid destruction.

Firstly, there is a risk of coercion. For, as I pointed out in 3.1, SMT demands the enhancement of someone who is insufficiently moral, immoral or amoral. But perhaps insufficiently moral persons would not want their morality to be enhanced, because they have less motivation to save the world. It is of course important to avoid the destruction of the world, but it is also important, especially for *morality* that every agent is autonomous. Even if we say that, according to J.S. Mill, harm to others can be a reason to restrict someone's freedom, what is permitted is not to modify their will, opinion, disposition or character, but only to restrict their *actions*. This is mainly because it does not further the greatest happiness of the greatest number, but there are also other reasons. For our will, disposition and character determine what is good and bad. Suppose we may freely modify our desire. Now we have two desires. One is to avoid the destruction, and the other is to avoid sacrificing ourselves. SMT told that we ought to modify the latter desire. But why should we not modify the former desire, when desire itself may be an object of modification? No sound ground has been given. So if we want to assess the value of the matter, we need some stable standard.[4] As well, if we want to avoid being rootless, we need to take our will, disposition and character as something given.[5] Hence, moral

4 This does not mean that the standard must be realistic or completely unchangeable. It may change, but it needs to be firm; at least something on which we can stand.

5 Cf. Sato [2011].

enhancement has the risk of unacceptable coercion and violation of autonomy of will.

Second, there is a risk of abuse. Powerful technology is easily abused by powerful groups. Particularly in the case of SMT, there is a structure that holds that sufficiently "moral" people enhance the morality of insufficiently moral people. But it is obviously difficult to decide who these sufficiently "moral" people are that never abuse biotechnology.[6]

Thirdly, there is difficulty in predicating or assessing the consequences. Global change occurs gradually over a long time, thus it may require many decades to realize moral enhancement. Stronger interference needs stronger reasons and justification, though current situations do not seem to provide a trustful prediction about what happens if we all become "moral".

Fourthly, to achieve the aims of SMT, we need to distribute drugs to a sufficient number of people, since it requires global collaboration. But it could exact enormous costs both financial and psychological. Many people psychologically resist taking drugs to remake themselves.

These four concerns seem difficult to overcome, behooving us to take a careful approach to moral enhancement. Indeed, Persson and Savulescu admit that "moral bioenhancement worthy of the name is practically impossible at present and might remain so far off for so long that we may not be able to master it, nor succeed in applying it on a sufficient scale, in time to help us to deal with the catastrophic problems" (Persson & Savulescu [2012] p.123).

However, it may also be true that such concerns are not as definitive as the researchers think. If the harm and probability are precisely predicted, and if their weight in fact overwhelmingly overcomes the harm related to the above concerns, it might be permitted to enhance the morality of insufficiently moral people even in a coercive way. There is, however, still a problem that this possibility does not show *moral* acceptability of moral enhancement in the end. This only suggests that moral enhancement is at least a *rational means* to an end, that is, to survive. Morality itself might surely be explained in terms of survival from the point of view of the theory of evolution. But still the question of whether something is good or not in terms of morality is obviously not similar to the question of whether something is good or not in terms of survival or rationality. And later, the question often depends on the answer to a former question. If we have a different conception of morality, for example, if we think that just sitting and accepting the destruction peacefully is a truly virtuous attitude, or if we regard that those who survive a tragedy through radical moral enhancement are not identical to "us" but some kind of transformed being, and if we find that the prescription to sacrifice our freedom for them is too demanding, then instrumental rationality will not require moral enhancement; on these conceptions, the rational thing to do might be to accept tragedy, or to look for other means which directly help us or do not demand that we sacrifice our freedom. These conceptions could be reasonable, so we need some grounds to dismiss them. So it is still open to question whether moral enhancement is morally and rationally acceptable. Its acceptability crucially depends on the normative theory on which

6 Cf. Harris [2011].

people stand, and it is not simple consequentialism alone like proponents of moral enhancement suppose.[7]

4.2 Weak Moral Enhancement Thesis and Private reasons

Now I shall turn to WMT. It is also apparently plausible, similar to SMT, yet it seems to lack advocates. I think this comes from the insufficient understanding of the morality the proponents of moral enhancement presuppose. Returning to the citation from Douglas, he said, "there are some emotions [...] whose attenuation would sometimes count as a moral enhancement regardless of which plausible moral and psychological theories one accepted"(Douglas [2011] p.470). This is true, and no moral theory would deem racial discrimination a good practice. However, each moral theory is different from one another in the way it rationalizes *why* discrimination is bad. Deontologists disapprove of discrimination because it violates our duty. Utilitarian disapprove of it because it reduces our total happiness. A virtue ethicist denies it because it is not what virtuous persons characteristically do. So here we must ask not only whether moral enhancement is good, but also why it is good.

Douglas, even after acknowledging that why Smith has reasons to enhance his morality is open to question, suggests that moral enhancement is acceptable, because (1) it brings at least one good consequence or (2) it has some intrinsic property such as the property of being an act of self-improvement (ibid. pp. 473-474). He considers first reasons as consequential and second reasons as non-consequential. In the following section, we will see that neither reason provides sufficient support for WMT.

4.2.1 Belief-Desire model and Good Relationships

Firstly, we will consider the consequential reason; that is, whether the consequence of moral enhancement really tends to be a morally good reason to conduct it. Initially, I will try to depict the scheme which proponents of WMT typically advance. In my view, proponents of moral enhancement focus on moral motivation, based on the "belief-desire model" of motivation theory. On the belief-desire model, motivation for any action consists of belief and desire. For example, if someone has a desire to want to treat people equally and a belief that to care for some members who belong to a certain racial group leads to treating people equally, then he must be motivated to care for them.

Now proponents of WME typically characterize the person who is permitted to enhance their morality as lacking moral motivation. Suppose Smith discriminates against his neighbor, Jones, because he belongs to a certain racial group. Here, according to the belief-desire model there are two possible accounts of Smith's discrimination. Smith fails to act morally, because he lacks (a) desire, to want to treat people equally or (b) a belief that to care for Jones (he must belong to a certain racial group) leads to treating people equally. Then, Smith lacks (c) motivation to

7 Or maybe, it requires a foundation provided by metaethics and ontology about morality. For example, H. Jonas, in his book *The Imperative of Responsibility*, develops his prominent arguments on the permanence of human kind based on his incisive theory of ontology about values, ends and being. cf. Jonas 1984.

care for Jones, and hence he discriminates against Jones.

The fact is that proponents (even most opponents) of enhancement presuppose the belief-desire model can be easily confirmed by some arguments. For example, Douglas holds that Smith lacks some *desire* (or emotion) to treat people equally, so he has a reason to enhance his *desire* (or emotion) by moral enhancement (Douglas 2011). On the other hand, J. Harris says that Smith lacks the correct *belief* or has a wrong one, so he has a reason to enhance his ability to have the correct belief using cognitive enhancement (Harris 2011). To put it another way, if someone has morally good desires and beliefs, and acts according to them, his action can be acknowledged as a morally good action. Whether one's action is good or not is assessed in terms of the desires and beliefs he has. Similarly, whether moral enhancement is good or not is determined by whether it can bring good desires or beliefs.

It is true, in some cases, that *rational* acceptability of moral enhancement could be decided from these aspects; that is, whether moral enhancement can instill the right motivation for moral action for the agent or not. If he wants to do morally good things, but does not have a desire to do his utmost to achieve it, he must be irrational in terms of instrumental rationality. In such a case it is rational to use moral enhancement to create a desire to treat people equally. However, whether moral enhancement could be *morally* acceptable or not, does not seem to solely depend on this point. In other words, moral reasons for accepting some actions are generally *relational*.[8]

Proponents of WMT think that the problem of the acceptability of moral enhancement depends only on the side of the agent. In other words, they consider whether an action is better or not in terms of the internal property of an agent's desire or emotion. For example, they think that aggression is bad, kindness is good and fairness is good. However, moral reasons for accepting some actions are generally *relational* and constructed by interaction with everyone who will be affected by the action.[9] Essentially, moral enhancement must be acceptable to all parties influenced by it; in this case, not only Smith, but Jones and all others discriminated against. Proponents of moral enhancement do not sufficiently consider the voices of victims. In other words, they ignore the elements of "reactive attitudes" in moral thought. They include an expectation of and demand for certain conduct from one another as rational and free agents.[10] We must carefully take the reaction of the addressee into consideration if our action is addressed for an agent (this is different from the case of a simple object, like a stone).

Now previous arguments on moral enhancement typically show the lack of sensitivity to victims by the analogy that compares moral enhancement to a treatment of alcoholism or arachnophobia (Spence 2008, Harris 2011, Douglas 2011b). They argue that someone who believes that drinking alcohol is bad but has no desire to stop drinking has a reason to enhance their desire to quit. But though

8 A point often stressed by the care ethicists. cf. Noddings [1984].

9 In this paper, the supposed relationship is a narrow personal one, but it can be wider. For example, H. Ehni and D. Aurenque point out that proponents of moral enhancement do not sufficiently consider social aspects of moral enhancement (Ehni & Aurenque 2012).

10 Cf. Strawson [1968], Darwall [2006].

alcohol does not complain and spiders also do not have any claim (perhaps), persons can have a claim. In our case, Jones or people being discriminated against might not be happy with artificially enhanced benevolence. He might see Smith's kindness as "varnish". He might think it is better that Smith honestly confesses his aversion. If Jones knows that pills produce Smith's kindness and that he hates Jones at heart, their relationship will not improve. Moral enhancement is rather likely to endanger any potential for a possible good relationship.[11] Smith's enhanced motive is better only for him, but not for Jones. So it is true that this moral enhancement brings at least one good consequence (for him), but it is still not sufficient for providing good moral reasons to enhance his morality. In the end, private reasons to be moral are not plausible.

So far we have seen that generally supposed private reasons are not sufficient for enhancement of our morality. However, still proponents of WMT could say, "But, regardless of a healthy relationship, it is better that Jones is not seriously discriminated against by Smith. This is a better consequence not only for Smith, but also for Jones. So there is at least one reason to enhance Smith's morality".

This reason seems to be more plausible than before. So if this is right, it might not be necessary to ban *all* types of moral enhancement. But, indeed, this response is difficult to adopt for proponents of WMT. For the term "seriously discriminated against" seems to show that this is already not a case of enhancement, but of treatment. Suppose that Smith does have a very firm belief that to discriminate against Jones is wrong, but still seriously discriminates against him. Do we think Smith is a moral person at all, in this case? Don't we think he misunderstands the word "wrong", or he is a kind of amoralist or immoralist who has morally bad desires?[12] If he misuses the word, then what he needs is not a biomedical intervention, but rather a traditional education. If he is an amoralist, this is not a case of "enhancement" of his morality, which makes well into better, but a kind of treatment, which makes bad into well.[13] And we can straightforwardly admit that some treatments are morally acceptable. For example, it could be a good use of biomedical intervention if Smith takes pills when Smith is attacked by a sudden, strong and hard-to-resist impulse to want to injure Jones and the harm to Jones could outweigh the harm of restricting Smith's freedom and chance to be virtuous.

11 Proponents of WMT might say that Smith could or should hide the fact that he was taking the pill. Then Jones would look at all the people who went to the drugstore with skeptical eyes. This is another bad consequence of moral enhancement, all things considered.

12 But it is said that there is still another possible interpretation of Smith rather than to be an amoralist. That is, he may have a second order desire to have a desire to not discriminate against Jones. Because of this second order desire, he has at least a moral character, so biomedical intervention to him deserves moral enhancement. This interpretation seems, however, not to work well. For he usually has this second order desire, since he himself understands himself as a non-moral person. So the presence of the second order desire does not guarantee that he is a moral person.

13 This distinction does not presuppose that treatment is morally acceptable and enhancement is not. However, contrary to the treatment that obviously has an instrumental goodness that relieves pain and disadvantage, the goodness of enhancement often needs to be justified further by all things considered judgment.

But, in the end, the consequence of moral enhancement which ignores a good relationship and even virtue (which I shall argue in the next section) might likely make the consequence worse in most cases, since we generally highly value a good relationship with others, and virtue seems to be essential to constructing such a relationship.

4.2.2 Virtue and radical enhancement

Now I want to discuss non-consequential reasons. Douglas holds that moral enhancement has some intrinsic property such as an act of self-improvement. Both deontology and virtue theory, which are representatives of non-consequential normative theory, would admit that self-improvement is a morally important practice. Even if we, however, acknowledge the importance of self-improvement, in fact, it is still open to interpretation whether moral enhancement is a euphemism for self-improvement. Then, the problem is, as I mentioned in 2.2, whether there is any difference between natural self-improvement and moral enhancement. Proponents of enhancement clearly see no difference between them.

However, I think there is a difference, particularly for virtue ethicists, because when we consider whether any practice is really moral, self-improvement or not, we need to consider not only the particular practice itself, but what the person had been doing and what they will do, over time. In other words, when we assess someone's morality, we do not only consider the particular actions, their consequences, and momentary emotion or belief, instead we consider their *character* and what kind of person they are — whether they are virtuous or not. This also holds true for when we choose our neighbor and construct a trusting friendship. So, regarding virtue ethics, self-improvement needs to imply the improvement of one's character. According to R. Hursthouse, a prominent virtue ethicist, a virtuous agent "really commits to the value" and acts "from a fixed and permanent state" (Hursthouse [1999] p.136). Nevertheless, the target of moral enhancement is typically certain emotions. Then, moral enhancement provides a makeshift solution for a morally painful predicament. What is *morally* important is to commit to some values sincerely and to acquire a fixed state through one's own harsh experience over time. So moral enhancement is not only not a means to be moral, but it is also likely to deprive the person of a chance to become a truly virtuous person.

However, some proponents of moral enhancement may say that we will be able to modify character itself in the future. Even so, I think there are still two problems. Firstly, even when a modified character unto itself is stable, if we once acknowledge intervening on behalf of the character, it can become unstable, as a result. We could freely change our character as many times as we want. This means that "stable" or "firm" loses its importance for assessing someone's character. Trustfulness of our daily assessment of someone is, however, usually tightly connected to and confirmed by stability and firmness of his behavior based on his disposition and character. In a world in which someone known to be shy may become so aggressive tomorrow, we may ponder how we can make a really trustful friendship? Then, I think the acceptance of such a radical enhancement of character undermines our practice of assessing someone and establishing a good relationship.

Secondly, such radical modification of character violates the premise that moral enhancement is a kind of self-education, because our identity is partly constituted by continuity of our character and disposition. So the more radical enhancement becomes, the more our identity between enhanced self and non-enhanced self will be lost. So, it might be said that this radical enhancement is not *self*-education, but rather to make someone into another person. Obviously, there is a difference between a reason to educate oneself and to re-create someone, because self-education could be justified for private reasons, but re-creating a person, which cuts off his identity cannot be justified. So proponents of moral enhancement would need to provide further argument that justifies a reason to re-create people.[14]

Proponents of WMT who hold that moral enhancement is similar to self-education, encounter the following dilemma. On the one hand, if proponents of moral enhancement want to keep it inside the realm of self-education, it cannot be so radical. If moral enhancement is, however, not so radical but modest and temporal, it cannot enhance our morality as a virtue. On the other hand, when the intervention is really radical it cannot be called self-education. Either will result in a disadvantage for the WMT, which posits that there is a non-consequential reason to enhance one's morality.

5. Conclusion

In this paper, we classified the arguments for moral enhancement into strong moral enhancement thesis and weak moral enhancement thesis. Then, it was shown that both arguments for enhancement upholding these theses were not sufficient. In particular, there are risks of coercion and abuse, which contribute to difficulty when assessing consequences and dealing with enormous costs, as in the case of SMT.

The situation with WMT is more complex. At first, there are consequential reasons and non-consequential reasons to enhance our morality. But consequential reasons do not work well, because moral enhancement does not usually bring better consequences for victims of discrimination, for example. Or if it does bring better consequences both for agent and victim, it is often not a case of enhancement but of treatment. Non-consequential reasons also do not hold, because if proponents of WMT uphold them, they will face a difficult dilemma; either giving up the notion that moral enhancement is a kind of self-education or that biomedical intervention changes our virtue. If moral enhancement cannot be self-education, proponents of WMT cannot appeal to a private reason, since it is not a purely private matter. If biomedical intervention cannot change our virtue, it is not *moral* enhancement. In either case, WMT advocates need to provide better justification to engage in moral enhancement.

14 Proponents of moral enhancement may bring up global reasons. Then the focus of arguments will be sent back to the issues that we saw in section 4.1.

References

Darwall S., 2006, *The Second Person Standpoint*, Harvard University Press.

Douglas T., 2011a, "Moral Enhancement", (eds.) Savulescu J., Meulen R. and Kahane G., *Enhancing Human Capacities*, Oxford University Press, 2011. pp. 467-485.

—— 2011b, "Moral Enhancement Via Direct Emotion Modulation: A Reply to John Harris", in *Bioethics*. http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8519.2011.01919.x/pdf. Accessed on 2013.1.30.

Ehni H. and Aurenque D., 2012, "On Moral Enhancement from a Habermasian Perspective", *Cambridge Quarterly of Healthcare Ethics*, 21, pp. 223-234.

Harris J., 2011, "Moral Enhancement and Freedom", *Bioethics*, Volume 25 Number 2, pp. 102-111.

—— 2012, "Ethics is for Bad Guys!, Putting the "Moral" into Moral Enhancement", *Bioethics.* http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8519.2011.01946.x/pdf, Accessed on 2012.10.25.

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

Jonas H., 1984, *The Imperative of Responsibility*, The University of Chicago Press.

Kovacs G.L., Sarnyai Z., Szabo G., 1998 "Oxytocin and addiction: a review", *Psychoneuroendocrinology,* Nov; 23(8): 945-62.

Noddings N., 1984, *Caring: A Feminine Approach to Ethics and Moral Education*, University of California Press.

Persson I. & Savulescu J., 2011,"Unfit for the Future? Human Nature, Scientific Progress, and the Need for Moral Enhancement", (eds.) Savulescu J., Meulen R. and Kahane G., *Enhancing Human Capacities*, Oxford University Press, 2011. pp. 486-500.

—— 2012, *Unfit for the Future*, Oxford University Press.

Sato T., 2011, "Human Enhancement and Human Nature", in *Applied Ethics: Old Wine in New Bottles*, Center for Applied Ethics and Philosophy, pp. 40-52.

Sean A. Spence 2008 "Can pharmacology help enhance human morality?", *The British Journal of Psychiatry*, 193 pp.179-180.

Strawson P.F. 1968 "Freedom and Resentment", in his *Freedom and Resentment and other Essays*, 1974, pp.1-29.

Arguments in Professional Ethics Concerning Justice and Autonomy with Regard to the Ill, Injured and Disabled

Paul JEWELL and Evdokia KALAITZIDIS

1. Introduction

A common conception of a just society is one that protects its citizen's freedoms. Another conception of a just society is one in which the citizens are treated equally. If a community values both of these ideals, as contemporary democratic societies typically do, then the issue emerges concerning how to balance them. An emphasis on personal and political freedom will not produce economic equality, and vice versa. In practice, whatever balance a community tries to achieve will be the subject of debate, negotiation and compromise. Part of that debate will be about what services should be provided by the State as apposed to those that are an individual's responsibility or best left to the market place. A leaning towards justice as freedom tends to decrease the role of the State and conversely a leaning towards economic equality tends to increase it. In the contemporary world, social arrangements vary. Some societies have extensive welfare arrangements, others place more responsibility on the individual, whilst insurance and philanthropy also play a part. Whatever the arrangement any community settles on, the extent of government intervention, the level of funding and the role of insurance and charities remain a matter for continuing debate.

In this paper, we enter the debate with a focus on the State's role with regard to those of its citizens who live with disabilities or chronic illness. Such people clearly need services, which libertarians might suggest are best accessed through the market place. Egalitarians might respond that justice requires the intervention of the State. We argue that the issue is not as clearly divided as that. The services that are required are often professional services. An examination of professions reveals two important features. One is that professions are highly regulated by the State. Their functions are not determined merely by the market place or their individual customers. The second feature is that professions are a social arrangement that a community needs in order for its citizens to enjoy lives characterized by self-determination, freedom and fulfillment. This is particularly true for people living with disability or illness. If professional services are seen as akin to goods that are bought and sold, then a libertarian can persuasively argue that the market place is the proper venue. On the other hand, if professional services are seen as a community regulated means of facilitating people's practical freedoms, then it can be argued that the State has a role in their provision, an argument that even libertarians might view favorably.

2. The problem

In well ordered contemporary societies that value justice, we expect that social arrangements will work for most of us most of the time. We work, we earn, we pay taxes. We engage professionals when we need their advice. We expect that there will be doctors whose expertise can be relied upon if we are ill, that there will be schools staffed with knowledgeable teachers and courts presided over by fair judges. We require government to provide us with security, protect our freedom and assist those of us who cannot help themselves. These social arrangements rest on some shared assumptions and notions of justice, notably that we are, by and large, free, self-determining persons who respect each others' rights and independence and co-operate rationally and productively with each other.

Our notions of justice are challenged when the assumptions do not hold. What social arrangements should we put in place for people who are not independent, or not co-operative and productive? Some people have a congenital disability or chronic illness which compromises or even precludes their participation in the community, their self-determination, their production and earning capacity. At the same time, they have significant needs, be they health, daily care, social interaction or mobility. How might our conceptions of distributive justice respond to this challenge (Jewell 2010)?

We might construct a notion of justice based upon need. We might try to arrange society such that people most in need, such as the disabled and chronically ill, were provided with the support they required, and suffered as little disadvantage compared to the rest of the community as we could manage.

To do so, however, would be to significantly discount our respect for liberty, self determination and a distribution based on productivity and individual entitlement. A libertarian may well acknowledge that disability and illness are certainly unfortunate and result in dire needs, but they are nonetheless nobody's fault and thus lay no obligation to redistribute wealth.

So a needs based concept of distributive justice is in conflict with the libertarian concept. How could we choose between them, or at least negotiate some sort of balance? This question could be seen as part of the traditional social contract issue discussed by Hobbes ([1651] 1962), Locke ([1689] 1962), Rousseau ([1762] 1968), Hume ([1888] 1902), Rawls (1971) and Nozick (1974). How can we balance the benefits of social participation against the compromise of political freedom? In this paper, we focus that question on seeking a balance between the needs of vulnerable people and individuals' freedoms, particularly with regard to economic entitlements.

3. An argument from the original position

To address the issue, we could invoke Rawls, who proposed that we imagine an 'original position' in which people make decisions concerning what sort of society they are about to set up. A just society would be a society that conformed to the

sort of arrangements such people would agree upon (Jewell 1983). To prevent them simply arguing from self interest, they would need to be ignorant of the circumstances in which they will find themselves once the social arrangements are realized.

> For example, if a man knew he was wealthy, he might find it rational to advance the principle that various taxes for welfare measures be counted as unjust; if he knew that he was poor, he would most likely propose the counter principle. To represent the desired restrictions, one imagines a situation in which everyone is deprived of this sort of information. One excludes the knowledge of those contingencies which sets men at odds and allows them to be guided by their prejudices. (Rawls 1971, 19)

For our purposes, we need to imagine that people in the original position understand that one of the advantages of society is the availability of professional services. They are also aware that in the society they are setting up, some of their number will have disabilities and chronic illnesses with high levels of needs for professional services. They do not know, though, whether they themselves will enjoy good health or whether they will be one of those with high needs. In the original position, would they institute social arrangements where the state accepted responsibility to make professional services readily available to those with high health needs, or would they make the acquisition and payment for these services the responsibility of the individual?

Rawls thought it unlikely that people in the original position would choose a slave society, because of the risk they would find themselves slaves rather than masters. Following this reasoning, we could imagine people in the original position contemplating the realization that some of their number would find themselves living with disability or illness. Most of them would not, but each faces a significant statistical chance that they would. It is plausible that they would prefer ready access to professional services rather than being left to fend for themselves should they find themselves ill or disabled.

This focus on professional services is a departure from Rawls, who was considering the difference between rich and poor. Arguably, the difference between healthy and impaired is not merely about money, but is about compromises to individual self- determination and practical freedoms. We can further argue that the aim of professional services is to maximize the self-determination and practical freedom of the recipient.

To advance this argument we should first examine the role of the professional in advancing the autonomy, self-determination and practical freedoms of people living with disabilities or chronic illness. We can then draw some implications from this for discussions of social justice.

4. The impact of illness or disability on personal freedom

Becoming and being ill or disabled is to experience a transition from being a moral agent to being a patient. A consequence for a normally self-determining individual includes physical conditions that significantly compromise their practical self-determination. The individual's normal functioning and flourishing is inhibited and suppressed. Rather than actively engaging with their surroundings, they are constrained by biological factors such as damaged organs and their condition is accompanied by significant impediments to their overall well-being. The gap below the individual's normal functioning and the consequence of illness or disability is likely to significantly compromise a person's freedom and ability to make decisions.

These conditions provide compelling reasons for a health care professional to fill in the 'gap'. A common way to envisage the problem is to see the role of the professional as the provider of a service, which in the professional's view will advance the patient's welfare. An alternative is to envisage the professional as a facilitator of patient welfare by providing the patient with advice, by making a recommendation so that the patient can advance their own practical self-determination. Both approaches require there to be an understanding of the intellectual and organizational elements associated with the professions. In a community that functions well, how do the nature and condition of individuals as clients relate to the intellectual and organizational elements of the professions?

The nature and condition of individuals as clients is similar for all professions. Communities include a number of individuals who have unmet needs, typically because they lack the intellectual resources to meet their own needs. The recognition of this fact motivates them to seek out the services of professionals. Professionals are engaged by individuals to provide them with assistance, advice or to make a recommendation on the basis of their unmet needs in relation to that particular profession's body of knowledge.

On this account, all clients are similar in that they are impaired and need a professional's advice in order to restore or at least improve their self-determination. One person might have a deficit in money managing ability, and so engages an accountant. Another is subject to legal constraints and needs a lawyer. Others still, such as children, lack literacy and numeracy, and require teachers.

Sometimes the individual seeking advice may know exactly what it is they want (perhaps to rid themselves of physical pain) but they may not have the technical competence to achieve it. For instance, the fact that they have a fractured limb may well be obvious to them, but a detailed diagnosis and appropriate treatment such as reducing the fracture requires assistance from others. In modern societies it is widely accepted that surgeons are responsible for treating individuals with broken bones.

Although individuals who become clients are similar in that they seek information and technical competence, the methods employed by professions to carry out that work are determined by the individual needs of the particular client. The way that the services are delivered vary, as do the implicit contracts between recipient and professional, and the language used to describe that relationship.

They may be called clients, users, consumers or customers. Interestingly but not surprising, in medicine, clients are described as 'patients' (AMA 2004, WMA 2006, ICN 2008, IPFCC 2012, Hutchings & Rapport 2012).

In all of these cases the client's needs and wishes are at issue, and this situation determines the methods employed by professions and the nature of the services offered. In determining how the work of the profession is conducted, the welfare of the client is of central concern to both parties and there is a client expectation that professional practice will advance the welfare of the client. Overall, professions have a variety of approaches to balancing considerations of welfare and autonomy and so do individual professionals. For example, people speak of 'being under doctor's orders', but in contrast lawyers ask clients 'What are your instructions?' Despite this difference, professionals have and apply professional competence. An important feature in the relationship between professionals and clients is the distribution and allocation and of responsibility for decision making (Bayles 1989, 75).

5. Professional expertise

The intellectual foundation of a professional practitioner includes a conceptual understanding of their field of knowledge (Holm 2011). There are two categories of knowledge that combine to constitute the intellectual foundation associated with professions. One category is derived from theory and the other from practice. Professionals are possessors of specialised knowledge and their professional practice is dependent upon them combining both types of knowledge.

The nature and condition of clients plays an essential role in the development of a profession's body of specialized knowledge and skills. Each profession's body of knowledge and skills contains a coherent set of understandings about the client's condition. For example, health professionals refer to a germ theory of disease when developing infection control practices. So when doctors and nurses perform surgical hand scrubs they do so because they possess an understanding of germs causing infections. That body of knowledge develops over time and involves a coordinated effort of many individuals and on institutions such as universities (Plato 1953, Newman 1964). A profession's specialized body of knowledge is a result of communal, disciplined, intellectual effort. It is a body of knowledge that is accessed and transmitted conceptually, has universal meaning to members across the same profession, and has implications for practice.

Research also plays a major role in the development of professional knowledge and skills. Professions are equipped with a tried and tested body of knowledge generated by research. It is not surprising, therefore, that society and individuals assume that professions possess a high level of expertise. The professionals' knowledge of universal laws of nature is applied to the everyday experiences of the client.

While each profession offers expert knowledge of a specific set of phenomena pertaining to clients and their conditions, there may be other relevant factors contributing to a client's condition which professions may recognise and

acknowledge, but may not have the expertise to deal with. The scenario of a lawyer dealing with a defamation case may serve as an example. The lawyer knows the law, and considers the case in the light of that knowledge. There may be factors that are not strictly part of the case, but which could affect the client's welfare, but these may be beyond the lawyer's professional responsibility. An historical instance of this is the case of Oscar Wilde, whose insistence on bringing a defamation case led to his downfall. Wilde's defamation action triggered an investigation into his personal life and resulted in his imprisonment (Ellmann 1987). This example shows that there can often be factors outside the scope of a profession's practice that could have a bearing on clients and on their condition. It is not always possible for professions to avoid producing other associated problems when they practice — as can be seen in Wilde's case. The job of a professional is not to run the client's entire life, but to give advice, which is sourced in, and limited to their professional knowledge base. No professional can be held responsible for the client's entire life. It is reasonable, however, to expect from professions that their advice be based upon expertise.

6. Client welfare as client autonomy

A client seeking advice from a professional is also an essential source of relevant information about his or her own circumstances. This may not be obvious to the client until they receive professional assistance. With professional assistance, clients get a fuller, clearer understanding of their condition and how it relates to other areas of their life. To make this happen, each of the two parties — the professional and the client — has a crucial role to play.

The role of the client in the professional relationship is to enter into a dialogue whereupon information is shared with the professional. The clients, as bearer of values and particular circumstances, need to reveal their condition to the professional. Professional consultation becomes a process involving the client and the professional with whom they are exchanging information. Typically, the professional frames questions to the client, eliciting specific information about the client's circumstances (Beach 2012). The client then provides the professional with the raw material and data to work on. The professional applies abstract principles to this data, achieving a specific understanding, and making decisions based on this understanding. This process is important, for it allows the professional to make a recommendation based on an appropriate body of knowledge. It also allows the quality of that advice and the soundness of that recommendation to be informed by the client in question.

Clients generally expect the professional to provide them with sound advice and to make a considered recommendation.

The development of a high level of professional expertise is associated with the skills required to process previous data and material in the light of new data collected from the client. Professional knowledge is expanding because the consultation process requires an intellectual shift back to previous data and material, but this is then done in a new context. As novel opportunities appear in

the professional context, links with theories are constructed. Sometimes phenomena are presented that do not correlate with pre-existing theoretical knowledge, because minimal or even no research has ever been conducted into the area. These cases may then encourage further research and theory development into the new phenomenon.

Although professions need to be actively engaged with client-centred problems in order to exercise expertise, each encounter with a client is different enough from any other to justify a role for standards, peer regulation, peer review professional codes and government control of practices. These organisational elements become mechanisms by which professionals are shown to be accountable to the wider community they serve. The wider community recognises the need for accreditation of a particular profession by developing mechanisms by which to achieve it, such as a licence to practise. This is an important feature of modern, complex social arrangements. Overall, the differences in knowledge and skills between a professional and a client necessitate the professional's accountability as well as commitment to technical standards and to the welfare of the client. So professionals such as doctors, nurses and disability carers perform a highly regulated, proficient social function.

Expertise in a professional relationship between the client and the professional is not entirely restricted to the professional. An important distinction between the two parties is that that professionals do not have the expertise in a client's values and are not qualified to make value choices significantly affecting a client's life plans or style (Bayles 1989,71). Compromised client self-determination gives rise to specific obligations in the professional. Professional roles and responsibilities evolve in the context of the wider social arrangements, embodying and expressing community desires, concerns and expectations. The client's ability to lead a self-determined fulfilling life is compromised by their condition, and so they access the professional's expertise to ameliorate the constraints on practical autonomy that they are experiencing. The professional's expertise and concomitant licence to practise is validated and regulated by the community.

7. A summary of the social role of professions

One stereotype of a professional's role is paternalistic. The clients of the medical profession are 'patients' and in common parlance people speak of 'being under doctor's orders' (Jewell 2008). A counter stereotype sees the professional as a hireling, so a lawyer might ask, "What are your instructions?" of a client who is presumably autonomous, in charge, knows what he wants and how to get it. An analysis of professions demonstrates the inadequacy of these stereotypes. A profession is a social arrangement. It is constructed, regulated and constrained by society. The purpose of professions is to provide advice to people who lack expertise. A client is one who has a knowledge deficit, who consequently has compromised self-determination, and who thus needs a professional's expert assistance and advice.

8. Implications for social justice

If we accept the argument that professionals should set out to maximize clients' practical self-determination, what impact might this have on our notion of justice? Justice is, after all, difficult to define. It has to do with fairness, equal treatment, impartiality, entitlements and rights, but what do these terms mean in practice? If we say that people should be treated equally, does that mean they should be treated the same? That does not seem right. If we advocate for justice for people with chronic illness or disabilities, surely we are arguing that such people should be treated differently. Treating everybody in the same way regardless of their differences would rule out paying any attention to people's different needs or circumstances. But if we are to treat people differently, how can we be fair to everybody, recognize everybody's universal human rights and be impartial?

9. Domains of justice

The notion of justice is not only difficult to define, but also occurs in different contexts. In medical ethics, it is one of the four fundamental ethical principles, alongside autonomy, beneficence and non-maleficence (Beauchamp & Childress 2009). In the legal system that deals with crime, it usually means *retributive* justice. Society seeks retribution by punishing criminals, and newspapers report verdicts as justice being done. A third usage is *distributive* justice, which is about the fair allocation of resources. Debates about social arrangements, taxation and welfare services typically use notions of distributive justice.

These three domains of justice—retributive, distributive and professional—clearly deal with different issues. The professional domain has an immediate relevance for professional ethics in the provision of health services to people with chronic illnesses or disabilities. The distributive domain is also important for clients of those services. The retributive domain ought to be irrelevant, but sadly it is not. The disproportionate number of people with cognitive disabilities and mental illnesses in prison is a matter of grave ethical concern (Dowse, Baldry & Snoyman 2009).

10. The domain of professional ethics

In the domain of professional ethics, if justice means that the same rules apply to everyone without favoritism, and that professionals should be fair and impartial, then the implication is that professionals should treat all their clients the same (Fremgen 2009, 12). On the other hand, professions have been established to meet clients' needs and the health care system has been established to provide services to people with health care needs (Holm 2011, 4). Clients do not all have the same needs. How can a professional, in order to practice justice, treat all clients the

same and individual clients differently? The solution is the understanding that all clients have needs and that they are the same in this sense — that they are clients who have needs. People who do not have needs are not clients (or patients, as the medical profession prefers to call them). Professionals' treatment of clients should be determined by the clients' needs and not by other factors. The amount of professional attention and services a client receives should not be influenced, either positively or negatively, by other factors such as race, gender or the personal likes/dislikes of the professional. Discrimination on those grounds would be unjust. We could further argue that all clients' needs are the same in that they all have compromised practical self-determination and need professional assistance to maximize their self-determination (Kalaitzidis 2009).

Admittedly, when applied to chronic illness or disability, this presents daunting challenges. When a client approaches a tax accountant, the client is taken to be competent at life in general, even if somewhat deficient in knowledge of taxation law. In practice, a patient in hospital is in a worse position than an accountant's client. In hospital, everyday things like mobility, food, clothing and social life are impacted severely by hospital routines, by the illness or injury and by the medical interventions such as drugs. Even so, for most hospital patients, this is temporary. For people with chronic illness or disability, constraints on practical self-determination are both chronic and life-pervasive. It may be tempting for society to give up in the face of these challenges and permanently institutionalize such people, thus making it more convenient for professionals to provide for their medical needs, but not their self-determination. Nonetheless, contemporary societies are rejecting that model of service provision and are de-institutionalizing. This trend clearly has implications that move beyond the context of professional ethics and into the domain of distributive justice.

11. The domain of distributive justice

Blackburn, when defining distributive justice, acknowledges that,

> The problem is to lay down principles specifying the just distribution of benefits and burdens: the outcome in which everyone receives their due. A common basis is that persons should be treated equally unless reasons for inequality exist; after that, the problems include the kind of reasons that justify departing from equality, the role of the state in rectifying equality and the link between a distributive system and the maximization of well-being. (Blackburn 1994, 203)

It was noted at the beginning of this paper that in a well ordered society social arrangements work for most people most of the time. Typically, citizens are assumed to be free, self-determining persons who engage professionals when they need advice concerning how best to carry out their goals. They require government to provide them with security, protect their freedoms and assist those who cannot help

themselves. Clearly, professional services are especially required for those citizens whose lives are constrained by illness, injury or disability. But at the same time, these conditions compromise their practical self-determination and simultaneously present obstacles to accessing those professional services. So to take up Blackburn's question, what should be 'the role of the state in rectifying equality and the link between a distributive system and the maximization of well-being' particularly with regard to citizens with chronic illness or disability?

Rose, in a review of the role of government in disability services, refers to social justice principles that:

- ensure that all people, irrespective of race, sex, disability or financial status, have equal access to government programmes;
- ensure access to opportunities to assist people to live as equal citizens;
- provide services to people in a way that best meets their needs and respect their rights; provide people with the opportunity to complain if the support they receive is inadequate or unsatisfactory in some way (Rose 1986, 97).

Treating everyone equally regardless of race, sex, disability or financial status, and at the same time meeting their (extensive and chronic) health care needs leads us towards a significant role for the state. It can be argued that there are common needs, such as roads, sewerage and policing that are best managed, supplied and financed by government agencies. Anyone might fall victim to criminals and we all need security and public order, so policing is a community need and therefore the responsibility of government. Any of us may fall sick and the community as a whole is more prosperous if its members are generally healthy, so health, including hospitals, infectious diseases control and sewerage systems, are the responsibility of government. Any of us might be born with a disability, or acquire a disability through accident or illness, so provision for disability is the responsibility of government. Justice, by this argument, is the entitlement that all citizens have to the support of the community in meeting individual needs in security, health and disability. Injustice would result if some people in need were abandoned by the community. In terms of distributive justice, this does not require that everybody be treated equally, or that wealth be equalized. It does require that resources to meet specified individual need be communally funded.

12. Justice as freedom

Being sympathetic to and providing for others' needs may be laudable, but it is not what everybody sees as the essence of justice. Sandel argues that debates about justice

> [...] revolve around three ideas: maximizing welfare, respecting freedom and promoting virtue. Each of these ideas points to a different way of thinking about justice. (Sandel 2010, 5)

An alternative definition of justice focuses on respecting freedom rather than on the community distributing services according to need. Flew defines it as

> to allocate to each their own [... the] individualistic ideal of securing for all their [...] diverse entitlements [...] to be contrasted with the collectivist ideal of imposing an equality of outcome (Flew 1979, 188).

If freedom rather than need is the essence of justice, then people are entitled to what they have earned, produced or acquired through fair exchange (Nozick 1974). Taxation as a means of assisting certain people on the grounds that they have needs is not justified. Such taxation is forcibly taking money away from people who have earned it and giving it to people who have not earned it. To do so would be unjust. Certainly chronically ill people have needs. Everyone has needs, perhaps in health, perhaps in housing, perhaps in personal relationships. It is up to individuals to figure out their own needs and to be responsible for meeting them. It is certainly not the government's responsibility. The role of government is to provide national security, law and public order; that is: to protect individual liberty. A libertarian might recognize the disadvantages facing a disabled person, but would argue that since disability is typically not anybody's fault, its concomitant disadvantages are not anybody's responsibility. The disadvantages are unfortunate, but not unfair. A community comprises a number of individuals freely and productively co-operating with each other, not a mass of people being coerced into uniformity. It is the satisfaction of individuals that matters, not some perception of the common good such as a somewhat arbitrary focus on equal access to health care. People should be respected as autonomous, rational, self-determining persons. By this perception, liberty is sacrosanct. In this view, justice is achieved when people can live their own lives as they see fit, and have their liberty and earned entitlements protected.

13. An irresolvable dilemma

It may well be that every community that attempts to establish social arrangements that are informed by concepts of justice faces an irresolvable dilemma. To arrange a society that meets people's needs and promotes their individual freedoms appears impossible. It may be that this dilemma is embedded in human nature itself. We all want the benefits of community and we all want to be self-determining individuals. The best we can do with this problem is to figure out how we can get the most of one side of the equation whilst giving up as little of the other. As our circumstances differ from each other's, or even as they change over time, so our concepts of justice might vary. Those of us in great need may argue for a distribution of resources that favors the needy and tends towards equality. Those of us who are capable and confident may prefer to rely upon our own unconstrained efforts. Perhaps we must resign ourselves to a continuous and never ending debate and negotiated compromise between the two ideals.

14. Risk, Freedom and Insurance

Various arguments can be brought to these negotiations. Acknowledging the devastating impact of serious illness or disability and concomitant needs, a libertarian might suggest that prudent individuals should secure insurance against such eventuality. In this way individual responsibility is preserved and needs met. Against this it can be pointed out that someone who is born with a disability has not had the opportunity to make decisions about insurance. If a pedestrian is run down by a motorist resulting in paraplegia, the driver's insurance will be called upon to meet the pedestrian's needs. But someone confined to a wheelchair as a result of a genetic impairment is not similarly covered.

This anomaly might be resolved by universal insurance (Soldatic & Dowse 2012). But to be universal, such schemes must be compulsory, so they favor the needs side of the equation rather than individual freedom.

We have argued that Rawls' theory of justice can be brought to this debate (Rawls 1971). Two (unrelated) points are worth noting in passing. One is that Rawls comes from the libertarian tradition and the other is that he did not apply his theory to the issue of disability. It seems obvious though, that it can be so applied by imagining that we were setting up a society and debating what would social arrangements should be implemented. We each could not know what our own circumstances were going to be in this new society. We each do not know if we would be ill or healthy, rich or poor, robust or disabled. Rawls himself proposed that we would want to maximize freedom. We would not choose to set up a slave society, for example, because of the risk of finding ourselves slaves rather than slave owners. He also thought that the most disadvantaged should be considered when allocating resources. Given the constraints on practical freedom brought about by chronic illness or disability, plus the economic burdens, Rawls conception of justice can be used as an argument for meeting the needs of people with disabling conditions.

15. Justice as distribution of autonomy

We have argued that people set up social arrangements to meet their needs. These arrangements include social services and the availability of professional advice. When people seek professional advice or social services, they are, in effect, seeking to cope with constraints on their practical freedoms. A person's ability to manage his/her own affairs is enhanced by advice from an accountant, or a lawyer, or a medical doctor. It is also enhanced by social services such as education or law enforcement. Illness and disability bring with them significant and pervasive constraints on personal freedom. A society that values justice would therefore implement arrangements that ensure the accessibility of social and professional services to such people. This approach would be driven by a conception of justice as the satisfaction of needs *and* by a conception of justice as the maximization of freedom.

16. Conclusion

The usual conception of a professional is an expert who provides services that maximize well-being; a doctor prescribes medicine, an accountant manages your money. We have argued that actually a professional typically provides advice and recommendations that allow us to practically maintain our self-determination, to lead our lives as effective, autonomous persons. We need a professional's advice because we have a deficit in self-determination. For most of us, the deficit is temporary and limited so we can continue to operate autonomously in general. For people with long-term illnesses or disabilities, the deficit in self-determination is significant and pervasive. For all of us, a professional service is at least as much about the maximization of autonomy as it is about promoting well being.

Conflicting notions of justice present us with a dilemma. We want to live in a society that maximizes our well-being and provides for our needs. We also want to live in a society that respects us as autonomous persons and maximizes our freedom. We cannot maximize each others' freedom and simultaneously compel each other to hand over the resources required to meet needs, especially needs which are ongoing, such as catering for permanent injury, illness or disability. On the other hand, such conditions themselves present significant obstacles to autonomy. Injustice is, essentially, about restrictions to self-determination (Weierter 2011). So if we value autonomy, and we recognize that professional services is as much about autonomy as about well-being, then we have an argument from justice to communally provide access to services for people with long term illness or disability.

References

Australian Medical Association (2004) *Code of Ethics for the Australian Medical Association,* Canberra: AMA.

Bayles, M. (1989) *Professional Ethics,* Belmont CA: Wadsworth.

Beach, M. (2012) Can patients be given too much room to make own decisions? *Medical Ethics Advisor*. AHC Media L.L.C., Accessed 7 January 2013 www.highbeam.com/doc/1G1-291064203.html

Beauchamp, T & Childress, J. (2009) *Principles of Biomedical Ethics 6th Ed*, New York: Oxford University Press.

Blackburn, S. (1994) *Oxford Dictionary of Philosophy,* UK: Oxford University Press.

Dowse, L. Baldry, E & Snoyman, P. (2009) Disabling criminology: conceptualising the intersections of critical disability studies and critical criminology for people with mental health and cognitive disabilities in the criminal justice system. *Australian Journal of Human Rights*, Vol.15(1): 1-29.

Ellmann, R. (1987) *Oscar Wilde*, London: Hamish Hamilton.

Flew, A. (1979) *A Dictionary of Philosophy*, London UK: Macmillan.

Fremgen, B. (2009) *Medical Law and Ethics*, New Jersey USA: Pearson Prentice Hall.

Hobbes, T. (1962) *Leviathan,* UK: Collins.

Holm, S. (2011) Who should decide the content of professional ethics? *Journal of Applied Ethics and Philosophy*, Vol.3: 1-26.

Hume, D. ([1888], 1902) *Enquiries Concerning the Human Understanding and Concerning the Principles of Morals*, Oxford UK: Clarendon.

Hutchings, H & Rapport, F. (2012) Patient centered professionalism. *Patient Intelligence*, Vol.4: 11-21.

Institute for Client and Family Centered Care, Accessed 7 January 2013 www.ipfcc.org/

International Council of Nurses. (2008) *Informed Patients. Position statement,* Geneva: ICN.

Jewell, P. (1983) *By What Authority,* Australia: Amada.

Jewell, P. (October 2008) Distributing responsibility for decision making in medical ethics. *Australian Review of Public Affairs.*

Jewell, P. (2010) *Disability Ethics: A framework for practitioners, professionals and policy makers*, Australia: Common Ground.

Kalaitzidis, E. (2009) *Client centered nursing ethics: A reconstruction of professional ethics for professional nursing,* Germany: Lambert Academic Publishing.

Locke, J. (1962) *Two Treatises of Government*, UK: Cambridge.

Newman, J. ([1854] 1964) *The Idea of a University, Defined and Illustrated,* USA: Holt, Rinehart & Winston.

Nozick, R. (1974) *Anarchy, State and Utopia,* USA: Basic Books.

Plato. ([c370 BC] [1871],1953) *The Dialogues of Plato, Vol. 2,* London UK: Oxford University Press.

Rawls, J. (1971) *A Theory of Justice*, USA: Harvard University Press.

Rose, A. (1986) *Australian Law Reform Commission Review of the Disability Services Act* (CTH).

Rousseau, J. ([1762], 1968) *The Social Contract,* UK: Penguin.

Sandel, M. (2010) *Justice: What*'s *the Right Thing to Do?* USA: Farrar,

Straus & Groux.

Soldatic, K. & Dowse, L. (2012) Another step forward for the NDIS, but details still missing. *The Conversation* Accessed 7 January 2013, https://theconversation.edu.au/another-step-forward-for-the-ndis-but-details-still-missing-11231

Weierter, S. (2011) Self determination culture and ideology. *Australian Journal of Professional and Applied Ethics*, Vol.13 (2): 33-46.

World Medical Association International Code of Medical Ethics, as amended by the *World Medical Association International Code of Medical Ethics*, as amended by the 57th WMA General Assembly, Pilanesberg, South Africa, October 2006, Accessed 7 January 2013 www.wma.net/en/30publications/10policies/c8/

The Unbearable Lightness of Personal Identity: Messages from Bioethics[1,2]

Cheng-Chih TSAI

Concerning the relation between metaphysics and bioethics, there are two main theses. On the one hand, some maintains, as exemplified by Conee (1999) and discussed in Shoemaker (2007), that they are more or less irrelevant, and on the other hand, some suggests that bioethics can benefit from metaphysical considerations about human identity, in particular, DeGrazia (2005) claims that a distinction between numerical identity and narrative identity may help us resolve some central issues in bioethics.

Unlike Earl Conee, I am not in the least bit skeptical of the relevancy of metaphysics to bioethics, and would indeed be happy to see DeGrazia's approach developed into a project as ambitious as the "Metaphysical Basis of Bioethics".[3] However, in this paper, I claim that some key notions in metaphysics that moral philosophers have adopted to do their bioethical analysis are simply non-well-founded. As a result, at present not only is metaphysics unable to help us resolve bioethical issues, some of its key notions may even be shattered to pieces under real impacts from bioethics. So, instead of hoping that metaphysics provides a sound foundation for bioethics right away, it is more reasonable to expect that bioethics will deconstruct metaphysics first — even if it is metaphysics in the analytic tradition that we are talking about — before it can be reconstructed and then be of any help to bioethics.

We shall look at three bioethical issues, namely that concerns the definition of death, that concerns brain transplantation, and that concerns dementia and advance directives, and see what metaphysical lessons we can learn from them. It turns out that each of them leads us directly to a metaphysical notion that is shaky, if not self-contradictory. To be more precise, the metaphysical notions in question are the unrealistic notion of other minds, the question-begging notion of personal identity, and the ungrounded notion of personal rights, respectively.[4] And I claim that, as the *referent*, the *sense*, and the *force* of personal identity all evaporate under close scrutiny, a sensible solution to all these three problems is to simply deny the reality of personal identity, and, instead, to resort to agency and social norm to deal with bioethical problems.[5]

1 This research is supported in part by National Science Foundation of Taiwan. The grant numbers are NSC 101-2914-I-715-001-A1 and NSC 101-2410-H-715 -001.

2 The author would like to thank the audience of ICAE2012, in particular Professor Ruth Chadwick and Dr. Paul Jewell for very helpful comments.

3 Along the line of Dummett (1993), *The Logical Basis of Metaphysics*.

4 For a more general skeptical view on personal identity, see Parfit (1984).

5 It is analogous to the case of physics. So far as physical interaction is concerned, what matters is not the identity of an electron but the physically measurable effects that such a particle — or, more precisely, such a wave function — can produce.

1. *Definitions of Death*: How could we ever refer to other minds?

When John is dead, we sometimes say 'John is gone'. But, *what* is gone? Surely, we do not mean John's body is gone, nor do we suggest that John's identity is gone, for otherwise the sentence 'John is dead' itself makes no sense. A straightforward answer to this question is that John's life is gone. But this is not helpful either. After all, we still have to know how the loss of life is to be characterized and why John's death should be so characterized. Or, perhaps, John's death has to do with whether his consciousness, or mind, has gone forever? Yet, where exactly is one's mind? We will come back to this problem later.

Despite that we are not so sure about what the death of a person amounts to, it poses, until recent decades, no real threats to our practice of death determining. Before the advances of modern life-sustaining technology, the death of a man was thought to happen at a particular moment — a man dies at the moment that his heart stops beating, his brain stops working, and his body stops functioning as a whole. There was little need, if there was any, to pin down a definition of death. However, we have learned now that, with the help of ECMO[6], one's brain can be functioning when her natural heart is no longer beating; or, as the case of anencephaly shows, one's higher brain can have no activities, or even be not existing, while her heart is still beating. As a matter of fact, medical professionals in most countries have now developed their own medical criteria for the determination of death. Permanent heart failure, the irreversible cessation of brain stem function, and the loss of higher brain activities are among the few criteria of death that we often come by.

On the face of it, the practice of organ donation can take advantage of the different definitions of death[7] so as to justify the timely removal of a desired organ from a "dead" patient. However, in reality, people seem to have strong intuitions concerning death, which are not easy to compromise. As the Terri Schiavo case shows, given that, excelled by the 1968 report of the Ad Hoc Committee of the Harvard Medical School, most countries have already adopted some form of brain-dead definition of death, when it comes to practical cases most people still could not easily adapt to it. Clearly, there is a tacitly adopted notion of death that guides us to prefer one criterion of death over the other. For some, Terri is dead once her consciousness is gone forever, and for others, Terri is not dead so long as she still breathes. It is even possible that one holds incompatible notions of death for different persons at different situations without ever realizing it.

Indeed, a definition of death that all people can agree upon is not easy to come by. But, before we can compare different definitions of death and decide which one is the best, there is something that we should know more about, namely, how the predicate 'is dead' is used. According to the Fregean program, to see whether the predicate 'is dead' applies to a proper name, we should first know what both the referent of the predicate and the referent of the name are respectively, and then decide upon whether the latter is an element of the former. However, there is

6 Extra-Corporeal Membrane Oxygenation.

7 DeGrazia has nicely sum up the evolution of the criterion of death in DeGrazia (2005).

something problematic here. If the name 'John' refers to a 4-dimensional person in space-time, which I shall term a 'B-person',[8] then no matter what definition of death we adopt, 'is dead' simply does not apply to 'John', because the person referred to by it is always there, not to be affected by death. If, on the other hand, 'John' refers only to a time-slice of a B-person, which we shall term an 'A-person', determined by some utterance of 'John' in certain context, then we will encounter a problem analogous to the problem of Existence Statement introduced in Devitt (1999), namely, 'John is dead' is meaningful only when it is false, or, equivalently, if 'John is dead' is true then it is meaningless.

How do we interpret our daily use of 'John is dead' then? There are three points worth noting here. First, as the truth of 'John is dead' apparently depends on the time of utterance, the referent of 'John' or 'is dead' should be time-dependent as well. Second, what is dead is not the identity of the person, but some essential feature of him which ceases to obtain after a certain time. Third, the criterion of the death of a person can usually be specified by some conditions on some particular part of the person.[9]

Let us be more specific. In saying 'John is dead', we are not, for reasons shown above, directly in the realm of Fregean semantics where 'John' refers to some individual and 'is dead' denotes some property of it. Rather, 'John is dead' amounts to a sentence of the form 'John's x fails to meet the y-condition at t' for some x, y and t.[10] Here are some possible candidates for x: (i) heart, (ii) brain stem, and (iii) cerebrum. Each of them can be given a practicable criterion y. Furthermore, the tendency, in most countries, of shifting from (i) to (iii) in defining death reflects the general assumption that a man can be pronounced dead only when his state of consciousness is forever gone, while the loss of consciousness can be characterized by the cerebrum's failure to satisfy certain criterion y.

Apparently, not all deaths are consciousness-related. For example, it is unlikely that we would imagine that the death of a plant, or even that of a giant squid, has anything to do with consciousness. But so far as the death of a human being is concerned, this conscious-relatedness seems to be a fair assumption and I have no dispute with it. It is just that there is a serious technical problem for one to refer to the consciousness of a person. Indeed, everyone knows that he himself has an inner conscious life, or a *mind*, for short. But, this by itself does not mean that we can refer to other people's minds by simply uttering the phrase 'his mind' and imagining that we can refer to that "mind" the way we refer to our own. As a matter of fact, under the Fregean scheme of reference, 'John's mind' fails to refer to anything, regardless of whether we adopt the description theory of reference or the causal theory of reference. Therefore, defining the death of a person by setting criteria on

8 In philosophy of time, A-theory and B-theory roughly correspond to the tensed and tenseless theories of time respectively.

9 This is so even if it is an A-person that we are talking about. For instance, for a brain-dead account proponent, the referent of 'John' may become nonexistent after he is brain-dead, yet 'the brain of John' still refers to some portion of a time-slice of a B-person in the space-time, and it is the brain's ceasing to satisfy certain condition that accounts for his death.

10 Actually, it can be more complicated than this. But we won't get into the details in this paper.

his "mind" is not an option at all. For a definition of death to make sense, we have to know what its associated criteria actually mean, yet while the referents of 'heart', 'brain stem', and 'cerebrum' can be made precise, the referent of 'mind' cannot. So either a "mind" can be reduced to be higher brain activities, or we should find other realistic ways to characterize it.

One possible way is to identify the mind with the character of John, which is characterized by the outward behavior of John, in particular, his responses to various stimuli, and these behaviors or responses are all observable or measurable in principle.[11] Alternatively, we can generalize the use of the term 'behaviors' further to include all physical activities in all parts of John that can be detected microscopically. In that case, the sentence 'John is thinking to himself' will be understood as indicating that some activities in his brain are currently happening and these activities are measurable provided that we have the technology.

Now, by the requirement that all entities and properties be outwardly observable or physically measurable, we should render any unresponsive — macroscopically or microscopically — individual dead. However, by the very nature of consciousness, it is not within the reach of others.[12] So there seems to be no way of ruling out the possibility that an unresponsive, in all physical aspects, individual is still conscious, thus remains alive — unless the thesis that there is a physically measurable effect associated with every mental activity has been proven. However, the latter thesis is nonsense, because in saying it we have assumed that we have an access to the mind already.[13] In sum, the act of associating one's *mind* with one of his or her body parts, characteristic behaviors or physical activities is more of a stipulation defining what consciousness is than a scientific discovery of the nature of consciousness. The whole situation is illustrated in Figure 1.

11 In a sense this is similar to the situation in quantum mechanics, where the "position" of a particle is meaningless unless we make a measurement of it. So, 'the mind of John' makes no sense unless we can detect a difference that it makes to John's behavior.

12 Just as Nagel's 'What is it like to be a bat' and Zhuangzi's 'The happiness of fish' have illustrated. See Nagel (1974) and Watson (1962) respectively. In the latter case, when Huizi said 'You're not a fish — how do you know what fish enjoy?' Zhuangzi replied 'You're not me, so how do you know I don't know what fish enjoys?' In this sense, DeGrazia's "narrative identity" is a meaningless concept as it assumes that we can readily refer to something in other people's mind.

13 Note that, one may think about setting up a scientific project establishing a relationship between one's feelings — pain, sadness, happiness, anger etc. — with his brain activities, and then use those physical characteristics to characterize other people's feeling. But, there is still a fundamental barrier that we could not cross: How do we know that the relationship holds always? There is always a possibility that for a demented person the relationship changes dramatically. For instance, for a person who can expresses her feelings, brain states *A* and *B* correspond to happiness and pain respectively, but for a demented one, *A* may correspond to pain and *B* to happiness rather. But, of course, whether this is the case, we will never know.

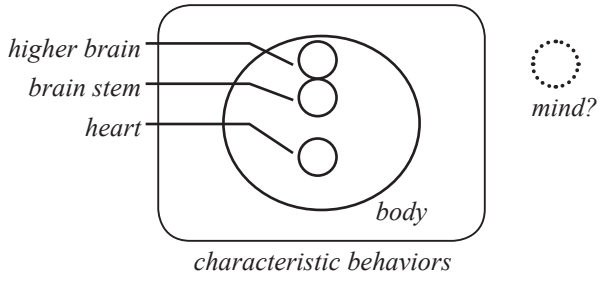

Figure 1

2. *Brain Transplantation*: Personal identity is to be created rather than discovered

Locke's classical example of Prince and Cobbler invites us to consider whether personal identity consists in spatial-temporal continuity or psychological continuity, while modern science reveals that the brain of a human being is responsible for much of his or her mental activities and hints that brain transplantation may not be as absurd an idea as it seemed to be. So, we can readily imagine the swapping of the brains of the prince and the cobbler here and see what we can say about the identities of the two after the swap?

I will consider, as a contemporary counterpart of Locke's Prince and Cobbler example, a brain transplant operation which removes the brains of Bush and Clinton and transplant them into each other's body. A natural question to ask here is that 'After the swap, which person is responsible for the affair Clinton had with Ms. L before the swap?' Apparently, it was the body of Clinton that did the deeds in question, yet it is the brain that is currently within (and controlling) the body of Bush that remembers the deeds. According to the spatial-temporal continuity theory of personal identity, it is the individual that currently has the body of Clinton that should be responsible for the scandal, while according to the psychological continuity theory it is the individual that currently has the body of Bush that should be responsible.

Recall that in the last section we have found that to refer to some person and apply certain predicates to his name, we need first to have some outwardly available characteristics that allow us to determine the referent of a proper name; furthermore, the fact that we seem to have a direct access to our minds does not entail that other people have similar minds and we can refer to them easily. Therefore, so far as this story is concerned, we should not simply assume that we have already known what 'Bush' and 'Clinton' refer to in the first place.

Fortunately, in this contemporary version of Locke's story, we do not need to worry about the impossibility of accessing one's psychological state: insofar as the psychological part is concerned we only need to look at the state or activities of the physical brain. Now, despite that I have used the second paragraph of this section to tell a story about the swap of the brains of Bush and Clinton, and the reader seem

to know what I mean by that also, it turns out that the story makes no sense at all unless we know what 'Bush' and 'Clinton' refer to, or at least how they refer, in the first place. In other words, in asking 'which one is Clinton after the swap?' one has committed the fallacy of Begging the Question, because in using the term 'Clinton', we have already assumed a grasping of the sense of the term in the first place. So there is no point to ask where *he* is after the swap.

To make it more precise, recall that here we have two individuals Bush and Clinton in question, and we talk about the body of Bush, the body of Clinton, the brain of Bush and the brain of Clinton. According to the unexamined notion of Clinton, there is a well-defined Clinton before the swap and 'the brain of Clinton' and 'the body of Clinton' refer to some physical objects in the world that we have no problem identifying even after the swap. But if we stick to the Fregean scheme of reference, then things simply do not work this way. We should know what Clinton is through space-time, before the phrase 'the body of Clinton' and 'the brain of Clinton' can make any sense. Yet, the referent of 'Clinton' in turn is determined by the referents of 'the body of Clinton' and 'the brain of Clinton'. How is this situation to be resolved?

I suggest that we pause and reflect on the basic process of reference. So far as personal identity is concerned, it is not that there is a pre-existent, well-conceived personal identity around us and we just need to choose a word to refer to it. Rather, the reference is in itself a process that shapes the individual. I shall emphasize, in contrast to Putnam's famous slogan 'Meanings just ain't in the head,'[14] that' *Identity simply is in the head"*. Or, more provocatively, "God created the world, and men created Adam."

To shape our idea of Clinton so as to pick out the referent of 'Clinton' in space-time, there are basically two ways.[15] Firstly, according to the description theory of reference, the reference of a proper name is achieved through a definite description, or a cluster of descriptions. This can be seen as being tied to the psychological continuity theory of personal identity, if the psychology is defined in terms of outward behaviors or responses of an entity that takes a human shape. Given that we have strong evidence that the brain is responsible for much of such outward characteristics, we can choose to characterize Clinton in terms of a brain. Secondly, according to the causal theory of reference, the reference of a proper name is achieved through a causal chain that links the name to its referent. Apparently, this is tied to the spatial-temporal continuity theory of personal identity. Given that the most direct causal link with a human being is through his body — we see his body, talk to his body and get struck by his body — we can, alternatively, choose to characterize Clinton in terms of a body.

14 Hilary Putnam, "Meaning and Reference" in Klemke (2000).

15 Recall that we have decided in the last section not to characterize an individual by his or her consciousness or internal mental state, as this would violate the requirement of realisticness. And this leaves us with two major strategies of characterization that correspond to the description theory of reference and the causal theory of reference respectively. And these two approaches in turn correspond nicely to the psychological continuity account of personal identity and the spatial-temporal continuity account of personal identity respectively.

I shall use 'Clinton $_{\text{qua-brain}}$' and 'Clinton $_{\text{qua-body}}$' respectively to refer to the two possible referents of 'Clinton' that are characterized by the two ways mentioned above. Note that without first specifying what notion of 'Clinton' we are adopting, 'the body of Clinton' and 'the brain of Clinton' are simply meaningless. Four "persons" that we can possibly refer to are illustrated in Figure 2 below.

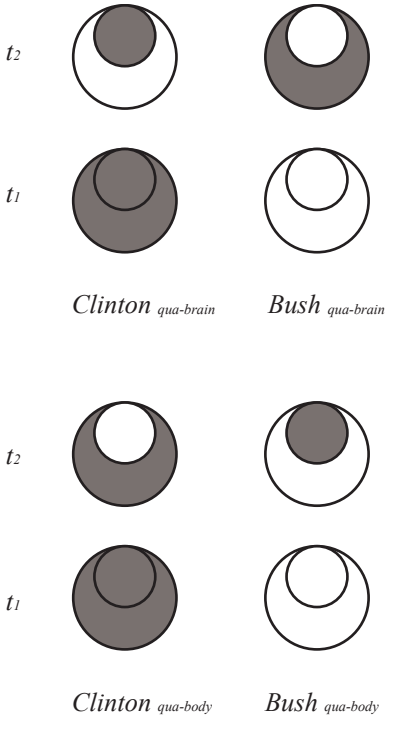

Figure 2

Here, the outer circles indicate the body and the inner circles indicate the brain. It is interesting to note also that, after the swap, the brain of Clinton $_{\text{qua-brain}}$ is in gray while the brain of Clinton $_{\text{qua-body}}$ is in white, and they are *different objects* in the world.

To identify Clinton $_{\text{qua-brain}}$, we look for his characteristic behaviors and responses governed by the brain in the body, and a description of what Clinton would normally do or respond is of great help for us to successfully refer to him; while to identify Clinton $_{\text{qua-body}}$, we look for the continuation of his original body, and a causal link of his physical form is of great help for us. So, in the case of Clinton $_{\text{qua-brain}}$, we would say something like 'despite that he now looks exactly like Bush, Clinton still has the mind of a Democrat' and in the case of Clinton $_{\text{qua-body}}$, we would say something like 'Fortunately, Clinton survived the lightning event, however his mind has experienced a dramatic change and he now becomes a Republican and behaves exactly like Bush.' In sum, what Clinton actually is, turns out to be a matter of how

we decide to see him as![16]

3. *Dementia and Advance Directives*: Personal identity itself is no ground for rights or responsibilities

Who is to decide the fate of a demented individual? 'The person himself' seems to be the right answer, but I think it is not. Firstly, in the case of dementia, the "person" is difficult to identify, as we shall see in a moment, and secondly, the identity is a matter of stipulation rather than a matter of fact, as mentioned in the preceding two sections. Furthermore, even if we have successfully identified a person, there is still no god-given rights for him to decide his future. What matters here is actually "Who is the boss?" — who has the power to exercise a right that can affect the fate of a human being — rather than "Who is the person?"

In the process of dementia, we do not see a sudden swap of brains as discussed in the last section,[17] but rather a gradual change of the "mental state", which, as I have stressed earlier, can be characterized by observable behaviors or measurable physical activities of the individual. Hence, on the face of it, there is no crisis of personal identity involved in the dementia case. However, in the literature, the Someone-Else Problem[18] has drawn our attention to the fact that the psychology of a demented individual may eventually turn out to be unrecognizable as being associated with the individual pre-dementia. That is, despite that a body at t_1 can be traced to a body at t_2, the brain inside the body at t_1 can be traced to the brain inside the body at t_2, and the two look the same, one would still be reluctant to admit that the person inside the body at t_1 is the same as the person inside the body at t_2, because the behaviors or responses of the two stages are so completely different.

This suggests that, unlike the case discussed in the preceding section, the continuity of a body plus the continuity of a brain is not sufficient to guarantee the sameness in personal identity. We need to look more closely into how the psychology of a person can be characterized here. So far as psychological trait is concerned, the continuity of psychology consists in two parts: the causal theory of reference seeks the continuity of a physical brain, while the description theory of reference seeks the similarity of characters characterized by the person's behaviors and responses. It is the dissimilarity of characters at time t_1 and time t_2 that drives the Someone-Else account proponents to suspect that there are actually two, or even

16 There is still another possibility that after the swap, one is very confused to such an extent that he or she can no longer identify Clinton and Bush. This is quite consistent with our position here — *identity simply is in the head*, and if your head cannot make sense of the identity of Clinton after certain event, then so be it. After John has been blown into a million pieces in bomb accident, there is no need for you to blame yourself for not being able to identify him any longer.

17 As discussed earlier, if we talk about the swap in terms of personal identities, then it begs the question, as there is *no* swap of personal identities at all. However, in terms of the actual movement of physical brains through space-time to join different bodies, the notion of "swapping" does make sense.

18 See DeGrazia (1999).

more, persons involved in a dementia case.

Now, we have two ways to take up this challenge. First, one can insist that as the dementia process is gradual and there is a similarity of mentality at the neighborhood of each moment, therefore there is simply no personal identity crisis here — there is only one person here, and it is just that sadly his mental state deteriorates to the extent that he could no longer recall things he has been through.[19] Second, one can claim that, in the case of dementia, even though we have no problem identifying a living being in the world, the different characteristic behaviors and responses of the being at t_1 and the being at t_2 suggest that we are dealing with two persons.

The latter approach identifies a person by his characteristic behaviors. However, it faces the serious problem of not accommodating changes of a character. Normally we can accept that John's behavior changes dramatically after some event, yet his personal identity is unaffected by the change. If we are to drop the continuity criterion of personal identity, and let our identification of persons be entirely based on descriptions of personality, then an essential ingredient of personal identity would have been lost. I shall, therefore, stick to the first option, that is, continuity condition should be carefully observed.

How then can we explain the fact that in the dementia case, we do often feel that the demented person, if *it* can be called a person at all, is not the same person as he used to be? In particular, in the case of advanced directives, while we seem to be obliged to follow the directive that the person has written down well before the dementia process begins, sometimes we are puzzled: when we try to remove the life-maintaining device, the person in the body seems to disagree with that directive and would beg us, with every possible means, not to do so. Since I have chosen not to adopt the Someone-Else account, I will maintain that it is the psychology of the *same* person that plays an essential role here and that the change of the psychology results in our getting two different instructions from the same individual at different stages of his life. People can change their interests, and it is the time-relative interests of the demented person rather than the very concept of personal identity that is in jeopardy here.

The following example further illustrates this point. If someone says to you on your eighteenth birthday that she would give you quarter a million dollars right away to grant you four enjoyable college years, provided that you would allow her to cut off your left little finger when you turn a hundred, it is likely that you would agree with her and receive the money, thinking to yourself that the you-at-a-hundred is simply too remote an entity for you to care about and that there is a good chance

19 This approach can be used to answer the worries of some philosophers who have spilt some ink on questions such as 'whether we have ever been a fetus'. There is no particular moment at which a pile of dust is formed, yet that doesn't consist in an objection to the identity of the pile. Similarly, I was gradually formed from dust and will decay gradually into dust. Which time best serves as the beginning or the end of me is a matter of convention or definition rather than a matter of fact. And normally we would not say that a person's identity terminates at a definite moment unless the person is blown to pieces in an instant by an atomic bomb, or goes through something like that. In particular, in case of dementia, most people would not think that the human identity ends at certain stage of the process.

that you would not even live to that age. Now, by the time that you are a hundred, it is very likely that you have long forgotten the promise you had made over eighty-two years ago,[20] and even if you still remember, you may not want to keep it. 'Why should my words eighty-two years ago be taken more seriously than my will today?' you might think to yourself. It is the younger you who made the promise with a young lady, received the money and spent it all, yet it's the poor and sick you-at-a-hundred who is about to lose a little finger to an old woman, how absurd![21]

This story shows that one can adopt different attitudes towards an event at different stages of his life and we can deal with such a situation without resorting to the change of personal identity. But, what are we to say about the rights or the responsibilities that each stage might have, if any, towards the fate of itself or of another stage? Some people no doubt would insist that since the two stages of you belong to the same person — you — the later part of you should be responsible for the deeds of the earlier part of you, and the earlier part of you has the right to prescribe in advance what should be done to the later part of you. But, why? There is NO metaphysical reason why this should be so. What we should be concerned about are practical questions like 'What stage of you is more relevant to the present situation?' 'Which agent is more capable of making a difference?' and 'What sort of "rights" and "responsibilities" should we ascribe to relevant agents so as to best guarantee social welfare in the long run?' etc... .

In particular, in the dementia case, there are at least three parties whose opinions concerning the fate of the demented individual we have to take into consideration. The situation can be illustrated as follows,

20 If you doubt that people would ever forget such a promise, just make the relevant ages to be 18 and 1000. If a human being can live that long, then there is a good chance that he would not remember a promise he made 982 years ago.

21 In the chapter 'Personal Identity' in Conee (2005), Sider imagines a suspect claiming that he is not the same person as the he who committed a crime several years ago. It urges us to reflect more on our criterion for personal identity and it assumes that the usual crime-and-punishment relation is based on the sameness in personal identity. But I think the latter assumption is not well-founded. Different stages of a person do often have different features and different time-relative-interests. But the very fact that these stages belong to the same person does not in itself say anything about the moral obligations between them. To explain the responsibility of the present you for the past deeds of you, we need to examine the genesis of rights and responsibility more closely. In particular, we should know that other people may demand something from you based on their beliefs that you-at-present is the same person as you-in-the-past, that they have means to punish you-at-present if you don't keep the promise made by you-in-the-past, and that such punishment of you-at-present will make them feel better. The most important thing is that your responsibility does not originate from your personal identity but from other people's wills and powers.

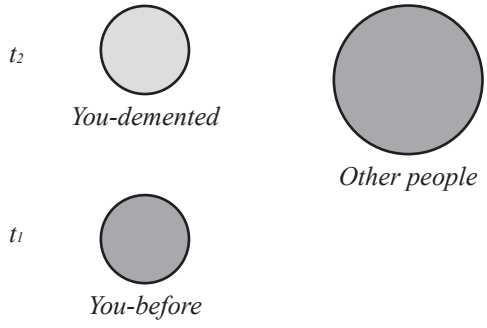

Figure 3

The You-before may have given an advance directive to remove the life-sustaining equipment when you are demented to a certain degree, while the You-demented may gesture that the contrary is to be done, and at the same time other people may express diverse opinions: your wife may claim that she is the only one who can represent you with a sane mind and, based on the fact that she has no affection to the present stage of you at all, ask the hospital to remove the life-sustaining device; your son may demand the same based on financial considerations, yet claim that his father has gone already and that the demented individual is but a breathing corpus of his father; and doctor A may be reluctant to take any action so as to avoid any future accusation, while doctor B refuses to do it because he believes that, as a doctor, he is to save people's life rather than to kill etc.. I do not claim to have an account that weighs all these opinions and settles the issue. I just want to stress that this is *not* a metaphysical problem concerning human identity at all. In a real-life bioethical crisis such as this, the fate of a person is to be decided by many people's — including his own — wills and powers, rather than by his opinion at some particular moment of his life alone. We need not grant a man who has written an advance directive any right that his directive shall be followed, if he simply has no way to defend it, either by himself, by someone else or by some social norm. Similarly, you have no responsibility to your past promise that allows the lady to cut off your left little finger when you turn a hundred, if there is no such lady to prosecute you based on her firm belief that punishing the later stage of your body would make her feel better about someone's not keeping the promise — a young man took her money and signed some contract and an old man who seems to be the spatial-temporal continuation of the young man refuses to let her cut off his left little fingers as stated in the contract.

4. Conclusion and applications

The metaphysics of personal identity does not help us resolve bioethical issues as readily as some authors have thought. Rather, bioethical issues invite metaphysicians

to reflect more upon identity-related concepts. We learn, among other things, three lessons in this paper: 1) Not only are other minds inaccessible to us, it is also meaningless for us to talk about them, unless we can characterize them in terms of outwardly observable patterns of behaviors or physically measurable events. 2) This characterization is more of a stipulation than a description of personal identity, and thus may vary from man to man. 3) Even if we have agreed upon a particular notion of personal identity, it in itself says nothing about personal rights or responsibilities.

This will free us from unnecessary worries about personal identities when dealing with bioethical issues and allow us to concentrate on finding out what relevant agents are involved in the situation, how they might exert their wills and powers, and what social norms can be established to maximize the overall welfare of the society and its citizens. The following two simple applications of this approach serve as an epilogue for this paper.

First, concerning the morality of abortion, we no longer need to 1) imagine how a fetus might feel about the world — there is no way to verify it; 2) decide whether a fetus should be counted as a person — it depends on how you define a "person"; and 3) worry about what rights are to be granted to it — no right is to be granted unless there is someone to fight for it.

Second, in Sinnott-Armstrong (2013), Walter Sinnott-Armstrong and Franklin G Miller propose that Dead Donor Rule may need to be lifted, and that the procurement of vital organs from universally and irreversibly disabled patients, which subsequently would lead to their death, is not morally wrong. The present paper provides a ground for their claim in the following senses. 1) It is unrealistic to imagine what those patients may feel towards the procurement especially when scientific evidences seem to suggest that chances are that they would not have an integrated mental experience of it. 2) We can decide not to identify our loved ones with the living organ that he or she used to inhabit, and learn to get used to this notion of personal identity. 3) By definition, a totally disabled person ceases to be an agent and has no way to act on his or her own, so any rights and responsibilities of his or hers can only be defended, demanded or imposed by others. In other words, if no one other than the totally disabled individual objects to the practice of organ procurement, there is nothing wrong to go on with it.

References

Conee, E. & Sider, T. (2005) *Riddles of Existence: A Guided Tour of Metaphysics*, Oxford: Oxford University Press.

Conee, E. (1999) Metaphysics and the Morality of Abortion, *Mind*, 108, pp. 619-646.

DeGrazia, D. (1999) Advance Directives, Dementia and 'the Someone Else Problem', *Bioethics* 13, 373-391.

DeGrazia, D. (2005) *Human Identity and Bioethics*, Cambridge: Cambridge University Press.

Devitt, M. and Sterelny, K. (1999) *Language and Reality: An Introduction to the Philosophy of Language,* 2nd edition, Cambridge, MA: MIT Press.

Dummett, M. (1991) *The Logical Basis of Metaphysics*, Cambridge, MA: Harvard University Press.

Klemke, E.D. and Geirsson, H. eds. (2000) *Contemporary Analytic and Linguistic Philosophies*, 2nd edition, Amherst, New York: Prometheus Books.

Nagel, T. (1974) What Is It Like to Be a Bat? *The Philosophical Review*, Vol. 83, no. 4, 435-450.

Parfit, D. (1984) *Reasons and Persons*. Oxford: Oxford University Press.

Shoemaker, D. W. (2007) Personal Identity and Practical Concerns, *Mind* 116: 316-357

Sinnott-Armstrong W. & Miller, F. G. (2013) What Makes Killing Wrong? *Journal of Medical Ethics* 39 (1): 3-7.

Watson, B. (1962) *Early Chinese Literature*. New York: Columbia University Press.

Rethinking Survivor Guilt: An Attempt at a Philosophical Interpretation

Satoshi FUKUMA

1. Introduction

In this paper I investigate the emotion of guilt, especially survivor guilt, to find its moral significance. The phenomenon of survivor guilt has been mainly studied in the fields of psychology from pathological and/or evolutionary viewpoints. After the Great East Japan Earthquake and tsunami (March 11, 2011) and the subsequent large-scale radiation leak that occurred at the Fukushima No. 1 (Dai-Ichi) nuclear power plant, many Japanese became depressed, suffered from fatigue, and then felt a vague sense of fear about their own future. One cause of this mood may be feelings of uneasiness, regret, sorrow, and unbearableness; that is, feelings of guilt about the victims of the disaster and the refugees who were evacuated from Fukushima Prefecture. In general, this feeling of guilt is called survivor guilt.

Survivor guilt is narrowly defined as guilt over surviving the death of a loved one, or broadly defined as guilt about being better off than others. The important feature of this guilt is that it is not directly connected with our voluntary wrong actions. Guilt, especially moral guilt, is an emotion one feels over disobeying a rule or command whose authority one accepts; in this case, the guilt is characterized as the feeling one has about one's wrongdoing. Typically, wrongdoing that is under one's control and for which one can take responsibility is an appropriate object of one's feeling of guilt. However, according to this characterization of guilt, survivor guilt seems to be an inappropriate and irrational emotion. This is because the person who feels survivor guilt has really not done anything wrong.

But is this feeling of guilt really inappropriate and irrational? If the people who suffered depression from survivor guilt are actually inflicted with an inappropriate and irrational emotion, should they instead be regarded as suffering from a pathological condition? Some psychologists (and also some philosophers) think so. However, I cannot agree with them because their understanding of survivor guilt is unilateral and insufficient. In this paper I try to investigate the phenomenon of survivor guilt from a philosophical perspective and explicate its features and practical meaning to justify this feeling of guilt in terms of the fact that only moral superiors can have it rationally and appropriately.

2. What is survivor guilt?: Psychological and evolutionary perspectives[1]

Survivor guilt is the type of guilt associated with the feeling of being better off than

1 To write this section I refer to O'Connor et al. (2000) and O'Connor et al. (2002).

others. In psychology, the concept of survivor guilt has both a narrow and broad meaning. This term was first used to describe the guilt that people may feel when literally surviving the death of another. Ch. Darwin described survivor guilt when he wrote:

> [...] under the sudden loss of a beloved person, one of the first and commonest thoughts which occurs, is that something more might have been done to save the lost one [...] in describing the behaviour of a girl at the sudden death of her father [...] she "went about the house wringing her hands like a creature demented, saying 'It was my fault'; 'I should never have left him'; 'If I had only sat up with him,' "[...] (Darwin 1872, 84f)

S. Freud described survivor guilt, referring to the guilt that he felt in the wake of his father's death. He noted "that tendency toward self-reproach which death invariably leaves among the survivors" (Freud 1985). W. Neiderland also studied survivor guilt among survivors of WW II prison camps who were found to be suffering from guilt, depression, anxiety, somatic symptoms, and sleep disturbances (Neiderland 1964).[2]

More recently, the use of the term survivor guilt has been expanded to include guilt about any advantage a person believes they have when compared with others, such as success, superior abilities, or a greater degree of health and well-being. A person may feel survivor guilt when someone close or beloved to him suffers misfortunes such as an illness or the loss of a job. Survivor guilt can even occur upon witnessing the suffering of strangers, such as reading about victims of violence, or seeing homeless beggars. So now survivor guilt has an extended meaning beyond the survivor's feeling after the death of persons familiar to them.

Survivor guilt may be a fundamental emotion that was developed by the evolutionary pressures due to living in small groups; it may promote social organization, insure an equitable distribution of resources, and prolong care for the young. Homo sapiens was psychologically adapted to life in hunter-gatherer cultures, many of which have been described as relatively egalitarian (Boehm 1993, 1997; Cosmides and Tooby 1992; Woodburn 1982). In an egalitarian setting, a high propensity for survivor guilt supported the behavior of sharing what is necessary for survival in complex group living. We may have inherited from our foraging ancestors an inclination to live in equality, and a tendency to experience survivor guilt in situations where inequity exists. From this perspective, survivor guilt has evolved along with various other forms of altruistic behavior and is a psychological mechanism that promotes sharing and concern for others.

For example, after the Great East Japan Earthquake, the feeling of survivor

2 William Niederland, a psychiatrist and a refugee from Nazi Germany, published a landmark study proclaiming the existence of a survivor syndrome. He listed a host of symptoms manifest in individuals who had survived Nazi persecution. They included chronic anxiety, fear of renewed persecution, depression, recurring nightmares, psychosomatic disorders, anhedonia (an inability to experience pleasure), social withdrawal, fatigue, hypochondria, an inability to concentrate, irritability, a hostile and mistrustful attitude toward the world, and a profound alteration of personal identity.

guilt may have caused people who suffered from the earthquake and tsunami but fortunately had some food and housing to share these things with the other victims. Additionally, most Japanese were willing to donate money and resources to the suffering people in the disaster areas, perhaps motivated by feeling of survivor guilt. From these examples, it appears that this feeling of guilt has positive and practical functions, and its mechanism can be explained naturalistically. Then one may say this sense of guilt is fitting because it is evolutionarily adaptive in the same way that the emotion of fear is.

However, evolutionary theory does not *justify* the phenomenon of survivor guilt. This is because when a person suffered from traumatic and extreme survivor guilt after the death of family members or friends in the Japanese disaster, he could not calm or eliminate this feeling of guilt even if he understood an evolution-based explanation of it. He would still feel responsible for their deaths and would experience feelings of self-condemnation or self-punishment, as well as severe depression. If that is the case, then would survivor guilt be an irrational and inappropriate emotion even if it is an innate emotion of human beings? In what follows, I will consider this question philosophically, focusing on the survivor guilt which survivors feel in the wake of the loss of their family and close friends (that is, a narrow definition of survivor guilt).

3. What does guilt tell us about survivor guilt?

To begin with, what kind of an emotion is guilt? Commonsensically, guilt is an uncomfortable feeling that one has because of having done wrong, causing anger and blame from others that is deserved. However, we may have feelings of guilt only because we *merely think* that we have done something wrong, despite the fact that we have not done anything wrong. It seems that there are two conditions in order for a feeling of guilt to be genuine or appropriate.[3] The first condition is: One really did harm or did something wrong to someone. The second condition is: The action is under one's control and one is responsible for it. These criteria for guilt, summarized by H. Morris' phrase "culpable responsibility for wrongdoing" (Morris 1987, 220), impose constraints on the application of this concept. According to this view, it is one's feeling of guilt, when one is actually innocent and so not guilty, that is inappropriate.

For instance, suppose that the tsunami rushed into a house where a mother and her two children waited for their father on March 11, 2011. When the father arrived at the high ground overlooking his home, he witnessed the house as it was swept up by the tsunami. He saw the fury of the tsunami, but he could only stand by doing nothing. His wife and children were drowned. He is now devastated by grief and an overwhelming sense of guilt. While it is understandable for him to feel grief for his lost family, why does he feel guilt? The man is not responsible for the tsunami or the death of his family. There is nothing he could have done, and he knows it. Logically and objectively, he understands that there was nothing he could have done.

3 In this paper I take "appropriate" to mean "rationally acceptable."

Yet, emotionally and subjectively, he is tormented by the thought that the death of his wife and children was caused by him.

His sense of guilt in this instance is survivor guilt. When a person feels guilt, he is disposed to recognize his action as a transgression of the legitimate claims of others and to expect them to resent his conduct and to penalize him in various ways. He also assumes that third parties will feel indignation towards him. Someone who feels guilty, then, is apprehensive about the resentment and indignation of others and the uncertainties which thereby arise. And he assumes that the victims or victims' family will be angry at him and blame him and his action. Subjectively, survivor guilt satisfies these requirements. When a person feels survivor guilt, he will have these dispositions and assumptions. However, in fact this person violated nobody's claims or rights, and he is not guilty. So nobody should blame him or his action. Without the blame from others, survivor guilt differs from normal moral guilt.

Then, would feeling survivor guilt based on the irrational belief of having done wrong to others also be inappropriate? Survivor guilt is the one type of guilt called "guilt without transgression."[4] Is guilt without transgression (or indirect guilt) an appropriate emotion? I think that there are three positions. The first is the view that survivor guilt is an inappropriate emotion (Rawls, Scanlon, and Wallace) (sec. 3); the second is the view that it is an appropriate but not moral emotion (Morris) (sec. 4); the third is the view that it is an appropriate and moral emotion (Greenspan) (sec. 5). Although I agree with Greenspan to a certain extent, I feel that her argument is still incomplete and would benefit from an additional component (meta-emotions based on appraisal theory and Watsuji's concepts of trust and guilt), which I will explain (sec. 6-7).

4. No, survivor guilt is not appropriate

J. Rawls, in *A Theory of Justice* (1999), explains the feeling of guilt at great length, and his view of it is called "the standard view" (Greenspan 1992, 287):

> He feels guilty because he has acted contrary to his sense of right and justice. By wrongly advancing his interests he has transgressed the rights of others, and his feelings of guilt will be more intense if he has ties of friendship and association to the injured parties. (Rawls 1999, 391)

For Rawls, the condition for which feeling guilty is morally appropriate involves one's transgressing the rights of others, not just doing harm or doing something wrong to others. Like Rawls, both T. Scanlon and R. Jay Wallace clarify this feeling of guilt as:

> [...] it is appropriate to feel guilt only when one believes that one has violated principles specifying what one owes to other people. (Scanlon

4 The "guilt without transgression" includes collective guilt, vicarious guilt, and existential guilt in addition to survivor guilt.

1998, 270)

> [...] one feels guilty for having violated an expectation that one holds oneself to, and so a particular occasion of guilt must be explained by the belief that one has in fact violated some such expectation. (Wallace 1994, 237)

Scanlon and Wallace think that this feeling of guilt is a reactive emotion, which arises from the *belief* or *judgment* that one violates "the rights of others" or "what one owes to other people" or "the other's expectation." This view is called a "judgmentalist" account of moral emotion (Greenspan 1992, 287). It presupposes that there is a causal or a constitutional connection between emotions and cognitive factors, such as judgments, beliefs, or thoughts. According to this view, if the judgment or belief that constitutes the feeling of guilt is inappropriate or irrational, then this feeling itself is also inappropriate or irrational. These three philosophers think that guilt is one of the moral emotions whose explanation requires the invocation of a moral concept and its associated principles.[5] Rawls says:

> In general, guilt, resentment, and indignation invoke the concept of right, whereas shame, contempt, and derision appeal to the concept of good. (Rawls 1999, 423)

Further, in analyzing the concept of guilt, Rawls distinguishes between *feeling* guilty and *being* guilty. He thinks that for a person to have a moral feeling, it is not necessary that everything asserted in his explanation be true. Additionally he remarks:

> It is sufficient that he accepts the explanation. Someone may be in error, then in thinking that he has taken more than his share. He may not be guilty. Nevertheless, he feels guilty since his explanation is of the right sort, and although mistaken, the beliefs he expresses are sincere. (Rawls 1999, 422)

Rawls assumes the possibility that one's *feeling* of guilt may not correspond to one's *being* guilty, but he does not mention explicitly whether such a non-correspondence feeling is appropriate or not. However, I think Rawls considers it as inappropriate or at least irrational because he suggests a rational feeling of guilt as follows:

> For if we suppose that, say, a rational feeling of guilt (that is, a feeling of guilt arising from applying the correct moral principles in the light of true or reasonable beliefs) implies a fault on my part, and that a greater feeling

5 "In general, it is a necessary feature of moral feelings, and part of what distinguishes them from the natural attitudes, that the person's explanation of his experience invokes a moral concept and its associated principles. His account of his feeling makes reference to an acknowledged right or wrong." (Rawls 1999, 421)

> of guilt implies a greater fault, then indeed breach of trust and the betray of friendships, and the like, are especially forbidden. (Rawls 1999, 416)

According to Rawls, for one's feeling of guilt to be rational, it needs a "fault on one's part" with based on "true or reasonable beliefs." Then if a person who thinks himself to be violating the rights of others but who in fact did nothing wrong (that is, he has no mistake or no culpability) has feelings of guilt, he has an irrational feeling of guilt even if his explanation of it is of the right sort and his beliefs about it are sincere.

In the judgmentalist account of moral emotion, guilt without transgression, and therefore survivor guilt, is not a genuine feeling of guilt. In this view, if a person is racked with survivor guilt, he seems to feel an irrational emotion and may be in a "pathological" state (Wallace 1994, Appendix 1). What he experiences would be the kinesthetic feelings and characteristic sensations similar to a genuine feeling of guilt, but it has no intentional or propositional objects (thing or fact of wrongdoing) that distinguish such genuine feeling. Experiencing feelings and sensations of this sort,[6] one might seek some transgression on one's part that would rationalize them (e.g., "If only I had not left home on that day, I could have saved them. What I did was wrong."). However, in fact, there is no such transgression.

5. Yes, survivor guilt is an appropriate, but nonmoral emotion

Contrary to Rawls and others, some philosophers consider survivor guilt to be an appropriate emotion; however, they think that it is nonmoral. For instance, Morris explains survivor guilt as follows:

> A person might be guilty just because of benefiting from a distribution that cannot be defended as fair or just or deserved. One's guilt would derive from being in an unjust position with regard to those with whom one identified. It would be this guilt that gives rise to impulses to redress imbalances. One's obligation would follow from one's guilt, not one's guilt from unfulfilled obligations. (Morris 1987, 236)

According to Morris, such a feeling of guilt would differ distinctly from moral guilt, for it would be a guilt that is independent of any choice to do wrong, and, as such, would be a blameless guilt. Then, by definition, it would be *nonmoral* guilt. Individuals who are *morally* guilty are viewed as justifiably condemned and justifiably punished for having set themselves apart from the community through insufficient attachment to its values (e.g., not harming others or respecting the rights of others). Nevertheless, in the case of survivor guilt, this practice associated with guilt is invalid.

However, Morris says that this feeling of guilt would manifest one's solidarity with others (ibid. 237) because it would derive from a moral posture toward others

6 A. Gibbard (2006, 200 fn10) calls these feelings and sensations the "flavor" of guilt.

and the world. When the relationship with those with whom one identifies is severed by a horrible accident, one feels guilty because of this *separation from* others. Such a feeling of guilt derives not from having done something wrong but from some conception of the moral solidarity of human beings (ibid. 232). The basis for this feeling of guilt is not a deed, but one's shared common humanity. From this perspective, Morris considers that although survivor guilt is not moral, it is "quite natural" (ibid. 234). He thinks that survivor guilt is a type of "guilt over unjust enrichment" that is by itself not pathological, unless it has an abnormal intensity or persistence, in which case it may then be pathological (ibid. 237).

Thus, Morris means that survivor guilt is a justified and appropriate emotion (ibid. 240). However, his view of "morality" is a little bit narrow because it only focuses on one's action or what one does.[7] When considering guilt and other basic emotions, I think that, *who I am*, in terms of my character and relationships and not just what I do, matters morally. Of course, character is expressed in action, but it is also expressed in emotions and attitudes (Sherman 2011). Moreover, many of the feelings that express character are not about what one has done or should have done, but rather about what one cares deeply about. Morality should be expanded to cover more than our actions and should include as moral emotion the guilt derived from a break with or a separation from the community of those about whom we care deeply.

6. Yes, survivor guilt is an appropriate and moral emotion

P. Greenspan examines survivor guilt on the basis of such an expanded view of morality. She classifies survivor guilt as one type of separation guilt[8] and rejects Morris' classification of separation guilt as necessarily nonmoral (Greenspan 1992, 301). She insists on detaching the grounds for guilt from the grounds for blame. According to her view, guilt without blame from others is a subjective but appropriate emotion. She calls her view of guilt a "nonjudgmentalist" view. It considers the subjectively guilty agent as feeling *as if* he were *morally* responsible.[9] In her own view, guilt amounts to discomfort with a certain evaluative propositional object and hence may be said to correspond to a judgment — though one can undergo the feeling without holding the judgment. In common practice, one can experience such a feeling — for example, when one's relationships with people with whom one has a sense of identity, or solidarity, or about whom one cares are abruptly ended by their death — and will not have to regard it as nonmoral. Therefore, this understanding — that one who experiences survivor guilt has the *as if* feeling or *construal* that one is morally responsible for a wrong (their death) and

7 While Morris admits the difference, his concept of moral guilt is based on that of legal guilt, and his concept focuses closely on one's action. See Morris (1988).

8 "Separation guilt" is characterized by the belief that one is harming one's parents or other loved ones by separating from them or by differing from them and thereby being disloyal (O'Connor et al. 1997).

9 This "as if" feeling is specified as "construal" by R. Roberts (1988), which might be called "appearances," or "seemings," or "quasi-beliefs" (Sinnott-Armstrong 2005).

therefore deserves punishment (at any rate, the emotional self-punishment of feeling of guilt)— seems to make more sense here in intuitive terms than Morris' claim that guilt in such cases is nonmoral. Feelings of guilt need not imply a strict cognitive acceptance of guilt. Perhaps *evaluative thoughts* are enough (Greenspan 1988, 43).

Wallace explains Greenspan's view of survivor guilt as:

> Persons who are subject to survivor guilt, for instance, need not really accept that they have done anything wrong in coming through a horrible accident alive. To account for cases of this kind, we should perhaps see guilt as connected not with beliefs about wrongdoing in the fullest sense of the term, but rather with our thoughts, which can structure our emotional experience even if we do not really accept them as true. (Wallace 2008, 108)

Thus, survivor guilt seems to be based on evaluative thoughts that are constituted of two elements: a sense of unjust enrichment and a sense of separation from others. For instance, a person may feel survivor guilt when he can get employment while his close friend cannot. He thinks that it is only luck that enables him to get a job and that he may not deserve the job more than his friend. In this case, his unjust or undeserved enrichment can cause a separation from his friend, either physically or mentally. However, in the case of a natural disaster like the Great East Japan Earthquake, the unjust or undeserved enrichment involves keeping one's *life*, or experiencing ("being donated") the *death* of relatives and close friends.[10] The person who could survive that disaster may presume that he could possess his life by another's death.[11] And in this case, the separation from others is bereavement. This separation cannot be repaired no matter what happens or what he does. (In the case of taking a job, the separation may cause the person to become estranged from his friend and only be repaired when the friend can get a comparable job.) Therefore, in this case his self-reproach and self-blame becomes more intense than when he feels normal moral guilt or precedent survivor guilt.

7. From the analysis of meta-emotions

In interpreting survivor guilt, Greenspan appeals to the "self-referential character of guilt" (Greenspan 1995, 165). In her view guilt functions as a kind of emotional self-punishment, and the failure to feel guilty counts as a possible object of guilt itself. (In certain situations, not feeling guilty in itself can be an object for which one should feel guilty.) She thinks that the self-referential character of guilt fits into the phenomenon of survivor guilt. This character can be more understandable from the analysis of meta-emotions.

10 See Ichinose (2011, the final chapter).

11 After the Amagasaki rail crash (April 25, 2005), some survivors of the disaster had such a presumption and faced physical and mental health problems. This crash increased my awareness of the phenomenon of survivor guilt.

After a major accident or disaster, not dying may be delightful and joyful for the survivors; however, the resulting separation from their loved ones can be painful and distressing for them. It seems that they have ambivalent emotions at the same time. C. Jäger and A. Bartsch (2006) contend that survivor guilt is an *intra*personal meta-emotion. In their definition, "meta-emotions" are higher-order emotions for lower-order emotions. For instance, when we have negative emotions for feeling malicious joy, our joy is an example of a lower-order emotion, while our negative emotions for this joy, because we construe it to be an inappropriate emotion in this context, are an example of a higher-order emotion. "Meta-emotions are elicited when a person appraises his or her own emotions in light of emotionally relevant appraisal criteria" (Jäger and Bartsch 2006, 194).

According to this view, survivors experience positive emotions about their fortune (such as relief and gratitude for living). But at the same time they feel guilty about *having these positive emotions* (ibid. 198). In the light of the tragic losses the event caused for others, the survivors feel — consciously or unconsciously — that any positive emotions related to the present situation are (normatively) inappropriate. According to this view, survivor guilt is not a special form of anxiety, but genuine guilt. However, it is not (the experience of) guilt over some "wrongdoing" in the ordinary sense, but (self-referential) guilt over being swamped with positive feelings despite the fact that feelings of grief and sorrow alone are perceived to be appropriate in the situation.

According to the appraisal theory of emotion, the meta-emotions we have toward the first-order emotion are either positive or negative, depending on our *perception* or *construal* of it as appropriate or inappropriate in the situation we are in. I consider that our evaluative perception that an emotion is appropriate or inappropriate in the present situation is not only normative but also moral. This is because the evaluation of its *appropriateness* must be based on the "social and moral propriety" of an emotion in those circumstances (ibid. 192). After a catastrophic event, survivor appears to control or restrain his positive emotion about his good luck (the first-order emotion) and come to feel guilty (the second-order emotion) for having such an emotion due to his sensitivity to or consideration for other sufferers. I think that his reflective and evaluative attitude about the first-order emotion and its appropriateness in the situation manifests his moral power and virtue. For survivor, unlike mourning over victims, feeling guilty about them and his own luck is not obligatory. Instead, it is a supererogation because some people may not feel guilty for having positive emotions about their survival luck and may feel satisfied with their fate. We cannot blame them for not feeling such guilt.[12] However, we would think that the people who can hold such feelings of guilt are more highly humane and sensitive and show more consideration for others. From these perspectives, survivor guilt is an appropriate and superior moral emotion.

12 If a person did not feel survivor guilt after his family or friends lost their lives in an accident, we would probably think less of him as a father or friend. However, we would neither condemn him for this nor force him to feel this emotion. Therefore, by this feature of survivor guilt, it is distinguished from the normal feeling of guilt.

8. From the Japanese ethical perspective

From the analysis of Jäger and Bartsch in the preceding section, survivor guilt is an *intra*personal meta-emotion; however, there would be a basis for an *inter*personal relationship that is the source of sensitivity to and consideration for others. What is this relationship that is necessary for one's feeling survivor guilt? In considering this question, I refer to Watsuji's philosophy. According to the anthropologist Ruth Benedict, Japanese culture can be classified as a "shame culture" whereas Western cultures are "guilt cultures." However, the Japanese ethicist Tetsuro Watsuji examined the concept of guilt in his masterpiece *Ethics (Rinrigaku)* (1937-49/1996). In what follows, I will examine his view of Japanese feelings of guilt.

The word *zaiseki*, which in Japanese is equivalent to the word "guilt," does not denote exactly the semantic meaning that the term "guilt" indicates in English (Watsuji 1996, 295). According to the English dictionary, the etymology of guilt is officially unknown, but it would have an origin that is similar to the German word *Schuld*, which means "debt" or "indebtedness."[13] The word *zaiseki* does not have these meanings. However, in Japanese, the words *sumanai* or *sumanakatta* are usually used to express an awareness of the badness of one's acts. (According to a translators' note [1996], the former word means "to have no excuse for and at the same time not to have finished doing," while the latter is the past tense form of *sumanai*.) Japanese makes use of these same words for expressing the concept of indebtedness, because the term *sumasu* means "to pay a debt" or "to pay one's bill." The term *sumanai*, which is the negative form of *sumasu*, precisely means indebtedness. Moreover, when what should be done is completely achieved, the Japanese term *sunda* (to have finished doing) is used. Therefore, Watsuji insists that the "Japanese can say that the word *sumanai*, which they quite ordinarily use to express an excuse, indicates something in common with the concept *Schuld"* (ibid. 296).

However, Watsuji also indicates a difference between *Schuld* (guilt) and *sumanai*. In Western culture, the meaning of indebtedness that *Schuld* (guilt) implies originally exists in the relationship (covenant) between God and human beings (they owe it to Him). On the other hand, in Japanese, the indebtedness that *sumanai* implies exists in the relations of trust within one's family, circle of friends, and society. If Japanese people inflict pain on others by their actions (e.g., violating an important commitment to a friend), they feel that they cannot be excused (i.e., *sumanai*) because they have betrayed the trust that others placed in them. Insofar as the awareness of badness consists in the consciousness of a betrayal of trust, there arises the feeling of *sumanai*, which establishes the significance of a debt or feeling of guilt (ibid. 296).

Then, for Westerners, the feeling of guilt is essentially thought of as a revolt

13 'Old English *gylt* "crime, sin, fault, fine," of unknown origin, though some suspect a connection to Old English *gieldan* "to pay for, debt," but OED editors find this "inadmissible phonologically."'(*Online Etymology Dictionary*, http://www.etymonline.com/index.php?allowed_in_frame=0&search=guilt&searchmode=none [accessed January 15, 2013].)

against God and His laws within an individual consciousness. Therefore, it is individualistic and *intra*personal (originating from the vertical relationship with God). In Japanese intuitions, this feeling is essentially and originally *inter*personal and is based on the terms among human beings (as in the Japanese word *aidagara*, which means the horizontal interhuman space or betweenness between humans). If Watsuji's analysis of guilt is correct (that the Japanese guilt feeling is interpersonal and community-based, originating from a betrayal of trust), could we cast new light on the issue of survivor guilt?

According to Watsuji's metaethical viewpoint, trust is the most basic component of and law regarding human relations. Not only our feeling of guilt, but also survivor guilt seems to have the moment which expresses the *betrayal of trusting relationship with others*. In section 5 I say that survivor guilt is based on evaluative thoughts that are constituted of two elements: a sense of unjust enrichment and a sense of separation from others. I can now add a third element, which is a sense of the betrayal of the trusting relationship with them.[14] It seems that the reason the survivor holds a feeling of guilt about his family or friends who were killed in the disaster is that: he feels that only his life was saved without any reason (sense of unjust enrichment) and that he was bereaved of them (sense of separation from others), and further he may think that he betrayed the trust of those about whom he cares deeply. Because he could not rescue them, he would assume that he did not respond to their trust that was conferred on him. As a result, recovering the world in which a trusting relationship is established — trusting in others and being trusted by them — is indispensable for the survivors in order to heal from their feeling of guilt. In Watsuji's concept of guilt, the trusting relationship between humans is an inducing factor for survivor guilt and is a necessary factor for recovering from it, so trust constitutes one of its fundamental elements.[15]

14 Though Rawls, whom I mentioned in section 3, discusses the relation between the feeling of guilt and trust, in his explanation, the essence of this feeling absolutely consists in transgressing the rights of others. He views a (betrayal of) trust as an additional element of this feeling (that is, if one has a trusting relationship with others whose rights one violates, one's feeling guilty is greater than one does not have). Additionally D. Velleman examines the connection of guilt and trust and interprets guilt as an anxiety about the loss of trust from others by having done something wrong (Velleman 2006, ch.7). His view also differs from that of Watsuji, who thinks that the betrayal of trust causes guilt. Beside this, Velleman analyzes survivor guilt and insists that this feeling of guilt would be rational if it were anxiety that a (fortunate) survivor has about the prospect being resented by the victim of a misfortune (ibid. 167).

15 It will be further necessary to perform a literature and empirical research to decide whether it is peculiar to the Japanese that "a betrayal of the trusting relationship with others" can constitute the third element of survivor guilt or it can hold true for people in other cultural areas.

9. Conclusion: A significance of philosophical interpretation of survivor guilt

From the discussion so far, I think that survivor guilt is an appropriate and rational moral emotion if the person does not want to eliminate it, prepares himself to retain it, and does not retrospectively regret having done so in the future, and moreover, has a *construing* of responsibility for the event that is counterfactual (with thoughts such as "I could have or should have done otherwise").[16]

H. Katchadourian claims that survivor guilt may have an "adaptive value" if it helps sustain an emotional engagement with those one has lost: "It gives a semblance of control over events that feel overwhelming and arbitrary — if we can do nothing else, we can at least feel bad about them" (Katchadourian 2009, 94). That is, by feeling guilt for the events that we could not cause or govern, we could get a *quasi*-sense of control over them. It is a way that we can impose a moral order on the chaos and awful randomness of a natural disaster's violence. It is a way that we can humanize the brutal disaster for victims and their bereaved families, and also for ourselves. If we cannot feel guilt about the event, we may then become as alienated from it as we are from its victims. Therefore, survivor guilt may also help preserve a sense of connectedness with those who have suffered. Painful as it may be, some people do not want to be rid of such guilt.

***

In our contemporary way of thinking, pathological or evolutionary explanations gradually occupy territory formerly governed by *moral* categories. Emotions, beliefs, and behaviors that used to be regarded as *morally good* or *bad* are now diagnosed as evolutionary adaptations or psychopathic disorders. These explanations have been beneficial and have expanded our knowledge about the world and ourselves, but by themselves, they are unilateral and insufficient. For instance, they cannot be appropriate explanations for issues involving the relationship between the deceased and us. I think that feeling concerned for the victims and guilty about their deaths is not only comforting for the bereaved families but also good for the deceased themselves. Commonsensically, it seems that it is better for the deceased that there is a person who cares about and moans for him rather than not. Although the deceased cannot experience the posthumous events relating to him, he can be changed posthumously by our attitude toward him and can undergo a good or wrong depending on the posthumous events.[17]

Therefore, the issues of survivor guilt are related to the metaphysics of death, so we have to take a holistic approach to dealing with them. Feeling survivor guilt

16 Though in fact he did nothing wrong, if he thinks factually of his assumed wrongdoing as "because I did not do it, she died," his feeling of guilt may be irrational and pathological.

17 I defend the view that the deceased can undergo a good or wrong by the posthumous events. See Fukuma (2009).

can serve a useful social function and express a moral virtue of subject and also is good for the deceased. Consequently, it is necessary to examine the issue of survivor guilt from more than one perspective.

References

Boehm, C. (1993) "Egalitarian Behaviour and Reverse Dominance Hierarchy," *Current Anthropology*, Vol. 34: 3, pp. 227-254.

—— (1997) "Impact of the Human Egalitarian Syndrome on Darwinian Selection Mechanics," *The American Naturalist*, Vol.150, No. S1, pp. S100-S121.

Darwin, C. (1872) *The Expression of the Emotions in Man and Animals*, London: John Murray. (http://darwin-online.org.uk/EditorialIntroductions/Freeman_TheExpressionoftheEmotions.html [accessed August 24, 2012].)

Cosmides, L. and Tooby, J. (1992) "Cognitive Adaptations for Social Exchange," in Barkow, J.H., Cosmides, L. and Tooby, J. (eds.) *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, New York: Oxford University Press.

Freud, S. (1985) *The Complete Letters of Sigmund Freud to Wilhelm Fleiss: 1887-1904,* transl. and ed. by J. M. Masson, Cambridge, Mass.: Belknap Press of Harvard University Press.

Fukuma, S. (2009) "Posthumous Harm or Posthumous Wrong?: A Philosophical Examination of Posthumous Events," Unpublished Manuscript (presented at The 4th International Tokyo Workshop on Applied Ethics and Philosophy, The University of Tokyo at Hongo).

Gibbard, A. (2006) "Moral Feelings and Moral Concepts," in Shafer-Landau, R. (ed.) *Oxford Studies in Metaethics*, Vol. 1, New York: Oxford University Press.

Greenspan, P.S. (1988) *Emotions and Reasons: An Inquiry into Emotional Justification*, New York: Routledge.

—— (1992) "Subjective Guilt and Responsibility," *Mind*, Vol.101, pp. 287-303.

—— (1995) *Practical Guilt: Moral Dilemmas, Emotions, and Social Norms*, New York: Oxford University Press.

Ichinose, M. (2011) *Ownership of Death: Philosophy towards the Death Penalty, Homicide, and the Use of Animals* (in Japanese), Tokyo: The University of Tokyo Press.

Jäger, C. and Bartsch, A. (2006) "Meta-Emotions," *Grazer Philosophische Studien*, Vol. 73, pp. 179-204.

Katchadourian, H. (2009) *Guilt: The Bite of Conscience*, Stanford: Stanford University Press.

Morris, H. (1987) "Nonmoral Guilt," in Schoeman, F. (ed.) *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*, Cambridge: Cambridge University Press.

—— (1988) "The Decline of Guilt," *Ethics*, Vol. 99: 1, pp. 68-76.

Neiderland, W. G. (1964) "Psychiatric Disorders among Persecution Victims: A Contribution to the Understanding of the Concentration Camp Pathology and Its Aftereffects," *Journal of Nervous and Mental Diseases*, Vol.139, pp. 458-474.

O'Connor, L. E., Berry, J. W., Weiss, J., Bush, M. & Sampson, H. (1997) "Interpersonal Guilt: The Development of a New Measure," *Journal of Clinical Psychology*, Vol. 53, pp.73-89.

O'Connor, L. E., Berry, J.W., Weiss, J., Schweitzer, D., and Sevier, M. (2000) "Survivor Guilt, Submissive Behaviour and Evolutionary Theory: The Down-Side of Winning in Social Comparison," *British Journal of Medical Psychology*, Vol.73: 4, pp. 519-530.

O'Connor, L. E., Berry, J. W., Weiss, J., and Gilbert, P. (2002) "Guilt, Fear, Submission, and Empathy in Depression," *Journal of Affective Disorders*, Vol. 71: 1, pp. 19-27.

Rawls, J. (1999) *A Theory of Justice*, rev. ed., Cambridge, Mass.: Belknap Press of Harvard University Press.

Roberts, R.C. (1988) "What an Emotion Is: A Sketch," *The Philosophical Review*, Vol. 97:2, pp. 183-209.

Scanlon, T. (1998) *What We Owe to Each Other*, Cambridge, Mass.: Belknap Press of Harvard University Press.

Sherman, N. (2011) JULY 3. "The Moral Logic of Survivor Guilt," *The New York Times*, July 3, http://opinionator.blogs.nytimes.com/2011/07/03/war-and-the-moral-logic-of-survivor-guilt/ [accessed February 9, 2012].

Sinnott-Armstrong, W. (2005) "You Ought to be Ashamed of Yourself (When You Violate an Imperfect Moral Obligation)," *Philosophical Issues*, Vol. 15: 1, pp. 193-208.

Velleman, J. D., (2006) *Self to Self: Selected Essays*, New York: Cambridge University Press.

Wallace, R. J. (1994) *Responsibility and the Moral Sentiments*, Cambridge, Mass.: Harvard University Press.

—— (2008) "Moral Psychology," in Jackson, F. and Smith, M. (eds.) *The Oxford Handbook of Contemporary Philosophy*, Oxford: Oxford University Press.

Watsuji, T. (1996) *Watsuji Tetsurô's Rinrigaku: Ethics in Japan*, trans. from the first half of *Rinrigaku* [Ethics] Vol. 1 by S.Yamamoto and R.E. Carter, Albany: State University of New York Press.

Woodburn, J. (1982) "Egalitarian Societies," *Man*, Vol. 17, pp. 431-451.

The Conflicting Terms of Environmental Justice: An Analysis of the Discourse

Morgan Chih-Tung HUANG

1. Introduction

There are three highly confusing notions within the environmental justice (EJ) literature — environmental racism, environmental equity, and environmental justice[1]. Most people have used these terms with little attention to how to define these concepts, nor to whether they are interchangeable in varied contexts. More than often, these terms have been considered synonyms and been conflated with one another.

Broadly, as it relates only to racial issues, environmental racism is the term used in the narrowest domain. Environmental equity, central to both income and racial issues, is at the meso scale. Together, these two terms are part of a larger environmental justice domain, with equations of EJ to environmental equity or EJ to environmental racism (sometimes even environmental racism to equity) to be found almost all the EJ literature.

As clear as it may be, these terms are used inconsistently and their domains are not clear. This article explores: What does the phrase environmental justice/equity/racism really mean? In answering this question, two foci are worth noting: how the concept of EJ has changed over time and how *science* is utilised in defining these terms. By using the tools supplied by STS (Science, Technology and Society), I attempt to demonstrate the process of constructing EJ in terms of terminology. At the end of this article, I will make clear how key EJ terms have changed in meaning and later became conceptually unstable.

2. Environmental racism: EJ in a racial sense

Reviewing the history of EJ, it is clear that EJ grew out of a series of anti-environmental-racism movements. The term "environmental racism" was therefore coined earlier than the other EJ terms. However, as the US government agencies have never adopted it to guide their policies, people continue to debate the meaning of environmental racism.

Despite the diversity of this term's interpretations, some consensuses can still be found. Benjamin Chavis, the former head of UCC (The United Church of Christ's Commission on Racial Justice), has frequently received credit for introducing this term:

1 It is notable that, here EJ is the overall phenomenon as well as (in environmental justice) a specific version of it.

> Environmental racism is racial discrimination in environmental policy-making and enforcement of regulations and laws, the deliberate targeting of communities of color for toxic waste facilities, the official sanctioning of the presence of life threatening poisons and pollutants for communities of color, and the history of excluding people of color from leadership of the environmental movement. (Chavis, 1994xii)

Another well-known figure defining the term is Bunyan Bryant. In his oft-cited book *Environmental Justice* (1995), environmental racism is defined as follows:

> It is an extension of racism. It refers to those institutional rules, regulations, and policies of government or corporate decisions that deliberately target certain communities for least desirable land uses, resulting in the disproportionate exposure of toxic and hazardous waste on communities based upon prescribed biological characteristics. Environmental racism is the unequal protection against toxic and hazardous waste exposure and the systematic exclusion of people of color from decisions affecting their communities. (B. I. Bryant, 1995:6)

With the statement of "deliberate targeting", the ongoing debate about this term's definition concentrates on the question of "intent". While both Chavis and Bryant suggest that accusations of environmental racism demand the proof of *intentional* discrimination, others emphasise that the presence of toxic waste in minority communities itself constitutes racism. As Bullard, also known as the father of EJ, has stressed (2000:98):

> *Environmental racism* refers to any policy, practice, or directive that differentially affects or disadvantages (whether intended or unintended) individuals, groups, or communities based on race or color. (Italics in original.)

The argument over intent is by no means trivial. In a typical EJ case, the industry and authorities completely deny the accusation that racial discrimination played a part in their decisions. In order to identify who is responsible, activists filed a series of lawsuits. In all these suits, the court insists that the plaintiffs are responsible for proving a sufficient pattern or practice of discrimination to support a finding of intent. Three court cases are reviewed here and these cases illustrate how legal terms, such as intent or discrimination, was considered in the manner of science.

2.1 Bean v. Southwestern Waste Management, Corp.[2]

In 1979, an African-American community started a fight against the siting of a solid waste landfill. Residents formed the Northeast Community Action Group

2 482 F. Supp. 673 (S.D. Tex. 1979), aff'd without opinion, 782 F.2d 1038 (5th Cir. 1986).

(the plaintiffs), NECAG[3], and filed a lawsuit to block the facility from being built. NECAG charged the Texas Department of Health (TDH) with discrimination in granting Southwestern Waste Management a permit in East Houston. This lawsuit was the first of its kind in challenging a siting decision on civil rights grounds. While the court acknowledged that this site was "unfortunate and insensitive", it still denied the plaintiffs motion as their statistical information failed to show sufficient evidence for defendant's discriminatory intent.

In *Bean*, the court based its analysis on demographic statistics about the minority populations. The court reasoned:

> The burden on [the plaintiffs] is to prove discriminatory purpose. That is, the plaintiffs must show not just the decision to grant the permit is objectionable, but that it is attributable to an intent to discriminate on the basis of race. (482 F. Supp. at 677)

To determine a pattern of discrimination, the court recognised that it is necessary to compare the racial composition in both "census tracts" and "the broader neighbourhoods" [4] where the facilities were located. The residents concentrated on two legal theories to establish this intent. First, they alleged the TDH's decision to issue the permit was part of a pattern/practice of discrimination. Second, considering the history of landfill siting and the issuance of permits, the plaintiffs contented that the TDH's approval of this permit constituted clear discrimination. Both legal theories were premised on large quantities of data and statistics and the plaintiffs relied on a series of statistical analysis on varying geographic areas in the proximity of the proposed facility.

In relation to the first theory, the court found that by 1978 TDH had approved seventeen sites. Of those, fourteen were sited in census tracts with a minority population making up to 50% of the community; ten were located in census tracts where 25% or less of the total population was minorities. In the "target area",[5] where 70% of their population was minorities[6], TDH had approved two sites for solid waste disposal. One was in a census tract having only 10% minorities; the other, the site being challenged here, was however located in a census tract with a 60% minority

3 According to Bullard (1999; 1995, 2000, 2001), at the urging of NECAG's attorney, Linda McKeever Bullard (also Bullard's wife), he conducted a case study of waste disposal practices in Houston. This is the start of Bullard's lifetime concerns on the EJ.

4 Please pay special attention to the units of analysis that the courts used in *Bean* and the following cases.

5 Target area is a term referring to an area designated by the federal government as low income. See: Rodriguez v. Barcelo, 358 F. Supp. 43, 45 (D.P.R. 1973). Since a target area was initially designed for identifying various income groups, the court recognised that the plaintiffs' definition and selection of "target area" is scientifically questionable. However, it asserted that this approach was somehow "useful" and worth to further examination. See: 482 F. Supp. at 677, 678.

6 Here, two scales were used to analyse its demographic features: target area and census tract. The plaintiffs' claims were based on target areas; the court however relied on census tracts in its analysis.

population. After reviewing the statistics above, the court concluded that because as many as half of the sites were located in an area where its minority population was less than 25%, statistical evidence alone failed to demonstrate a clear pattern or practice of discrimination.

In the other theory, the plaintiffs provided three sets of data to support their argumentation. The court however rejected all of them. The first set of data focused on the two solid waste sites in Houston. Since both of these sites are located in the target area, this proved discrimination, the plaintiffs argued. They further stressed the fact that the target area's population amounted to only 6.9% of Houston's total population; it however hosted "100%" of this city's type I landfills. The court used the same grounds, one of the two sites was in a White tract (less than 25% minority population), to reject this data set. Meanwhile, two sites did not constitute a statistically meaningful sample to infer discriminatory intent, the court noted.

The second data set concerned the total number of these facilities. The plaintiff's argument was that 6.9% of Houston's population lived in the target area; however, they hosted 15% of these facilities in the city. Since 70% of the target area population was composed of minorities, the plaintiffs contended that this disparity constituted discrimination. The court however stated that, outside the target area, the other facilities were located in 70% White tracts. Thus, no discrimination can be determined from this data either.

The final data set separated Houston into two halves. The eastern half contained 61.6% of the minority population and hosted 67.6% of this city's solid waste facilities. The western half however had a 73.4% White population, and only 32.4% of waste facilities. The plaintiffs alleged that the disparity (67.6% vs. 32.4%) constituted discrimination. Again, the court disagreed with this argument because it neglected the fact that the industrial area was located in the eastern half of the city. In other words, the alleged disparity may simply be the result of industrial agglomeration.

As the first lawsuit against environmental racism, *Bean* gave the EJ movement a very "scientific" start. *Bean* heavily relied on statistical evidence to support its case. In so doing, the legal terminology, discriminatory intent, was translated into scientific/geographic patterns between race and the location of facilities. According to the court, however, the available data, both city-wise and in the target area, were unable to prove that such a pattern really exists.

2.2 East Bibb Twiggs Neighborhood Association v. Macon-Bibb County Planning & Zoning Commission[7]

East Bibb Twigg (Bibb), Georgia, also challenged the local authority's (the Planning and zoning Commission's) permit approval by arguing racial discrimination. To evidence their claim, the plaintiffs constructed three theories and data sets. Having reviewed the administrative history of the permit and the data derived from both "census tracts" and "governmental districts", the court ruled that the plaintiffs had failed to demonstrate a clear discriminatory impact or intent.

To begin with, the plaintiffs pinpointed the discriminatory impacts of this

7 706 F. Supp.880 (M.D. Ga.) aff'd 896 F.2d 1264 (11th Cir. 1989).

permit decision. They argued that the proposed facility was located in a 60% Black census tract. As a result, issuing this permit would most impact the African American community. The court, however, noted that the only other landfill permitted by the Commission was located in a 76% White census tract; therefore no obvious patterns of discriminatory impacts could be found.

The plaintiffs then changed their unit of analysis and further argued that putting the one Black and one White landfills in a bigger picture, both of them were in factually a Black governmental district where the Black population were around 70%. By stressing the importance of using "census tracts analysis", the court however denied the plaintiffs' contention.

The plaintiffs' final theory relied on the history of racially based decisions by the Commission. They argued that fifteen years ago the Commission had issued a report recognising the existence of racial discrimination in this area. Furthermore, they questioned the reasons why the Commission changed its decision after initially denying the permit. To this argument, the court reviewed both the administrative history of this particular permit, and the past record of this Commission's previous decisions. It found that the Commission's prior permit decision was in a White census tract. And, no evidence showed that the Commission has suddenly changed its zoning classifications or relaxed its permit-issuing standard. The court ruled that the Commission had no history of tending to authorise facilities in Black communities.

Two issues are important in *Bibb*: the standards for determining a claim of discrimination, and the analytical unit used to determine the impacts of a site. Regarding the first, the court closely followed the test suggested in *Arlington Heights*[8]. *According to the US Supreme Court, in order to prove a claim of discrimination in violation of the Equal Protection Clause, a plaintiff must establish two things: the governmental officers acted with a" discriminatory intent" and their action had a" discriminatory impact". Under this standard, scientific evidence of a disproportionate risk distribution alone is not* enough to support a discrimination case. In terms of the second issue, the court emphasised the importance of using census tract analysis. That is, the court compared census tracts with and without facilities to determine the existence of the discriminatory impact and the role of historical discrimination. In sum, the *Bibb* court, as some pointed (Collin, 1992:526), almost overstated the importance of using census tract test; and therefore other relevant analytical units were overlooked.

2.3 R.I.S.E., Inc v. Kay[9]

In *R.I.S.E.* (Residents Involved in Saving the Environment), plaintiff challenged the siting decision of the county commissioners. Originally, *R.I.S.E.* was concerned with environmental problems, like noise, odors, and decreased property values, though after its initial opposition to the project failed, its focus shifted to environmental racism. In King and Queen County, Virginia, the population was approximately

8 A good summary see: *Bibb* at 884. More fully see: *Arlington Heights v. Metropolitan Housing Corp.*, 429 U.S. 252 (1977).

9 768 F. Supp. 1144 (E.D. Va. 1991).

42% Black and 57% White. After demographic analysis, the court confirmed a disproportionate impact on the Black community; thus, the plaintiffs had met the first step of the discriminatory equation set forth in *Arlington Heights*. Nonetheless, the court further noted the impact of a discriminatory action provides only a "starting point" for determining one's discriminatory action. The court then ruled that the plaintiff had failed to supply sufficient evidence to show the choice of site itself was intentionally discriminatory.

Unlike the previous cases, census tracts were *not* used in *R.I.S.E.*, because African Americans were extraordinarily concentrated in much smaller areas around the target landfills. The court was not dependent on a particular unit of analysis, like census tracts or zip code, but reliant on the concentration of African American at various *distances* from a particular facility. Within a half-mile radius of the targeted siting area, African Americans accounted for 64% of the population; specifically, there were 39 Blacks (64%) and 22 Whites (36%) and 61 people in total living in this area. Further, 21 Black families and 5 White ones lived along the 3.2 mile road leading to the proposed landfill. Investigating this particular landfill alone, there seemed to be a racially disproportionate distribution. In order to identify a disproportionate impact on the Black residents, the court insisted on examining the demographics and the siting procedures of the past four landfills.

The Mascot Landfill, sited in 1969, was the first one examined. The racial composition of the population who lived within a one-mile radius of this landfill was 100% Black. Moreover, there was an important Black community church only two miles away from this landfill. The second site, the Dahlgren Landfill, was sited in 1971. When the *R.I.S.E.* was filed, within a two mile area, 90-95% of the population was Black. Owenton Landfill was the third one discussed by the court. When it was first established in 1977, all residents within a half-mile radius of it were Black. The First Mount Olive Church, an African American church, was within a one mile radius of the site. The fourth site, King Land Landfill, was developed in 1986. This landfill was originally located in a predominately White neighbourhood. Although it had received a state permit initially, it was shut down when the county obtained an injunction to stop its operation, due to environmental violations and the community opposition.

While taking note that "[t]he placement of landfills in King and Queen County from 1969 to the present has had a disproportionate impact on [B]lack residents,"[10] the court found that "official action will not be held unconstitutional solely because it results in a racially disproportionate impact. Such action violates the Fourteenth Amendment's Equal Protection Clause only if it is intentionally discriminatory"[11],

To sum up, as a high proportion of minority populations were so close to the sites, unit of analysis was irrelevant in *R.I.S.E.*. However, *R.I.S.E.* begs a question of why/how a particular "distance" was chosen by the court. Moreover, according to the court, finding a disproportionate impact was only the first step to determine a racially discriminatory intent. After reviewing this county's siting procedures, the court concluded that there was nothing unusual because "the Board appears to

10 *Id*. at 1149.

11 *Id*. at 1149.

have balanced the economic, environmental, and cultural needs of the county in a responsible and conscientious manner."[12]

2.4 Pattern seeking: Scientifically addressing the harm of environmental racism

Having reviewed these cases, it is not hard to recognise why Bullard (2000:98) insists that intentional discrimination is *not* the most plausible way of identifying environmental racism. He asserts that, since the roots of racism may be so deep, discriminatory outcomes/impacts may not always result from discriminatory intent. By arguing that "harm perpetuated by benign inadvertence is as injurious as harm by purposeful intent" (cited in Ringquist, 2006:251), later commentators reinterpret environmental racism as "any" decision-making process and distributive patterns that results in unequal burden on the minorities. This new interpretation originally attempted to avoid the difficulty of proving someone's intent. The problem however is not yet completely solved as plaintiffs/activists still need proof of disproportionate impacts, but exactly what constitutes these impacts and how best to test/measure them remain in dispute.

So as to prove a defendant's discriminatory behaviour, plaintiffs have to objectively illustrate the pattern/impact of disproportionate siting-decisions. In other words, legal terms like discrimination and intent are re-conceptualised as a scientific/statistic terms in the analysis of a right violation. The court uses scientific data, as the first step, in determining whether the plaintiffs, as individuals, have been denied rights shared by others on the basis of group membership. Some commentators (such as Foster 1993) describe this approach as a (constitutional) civil rights paradigm, in which the harm of environmental racism is defined as either the denial of a right to a clean environment or the right not to shoulder an unequal burden of toxins.

There are three notable features concerning civil rights paradigm. Firstly, racism in a legal sense is construed to mean conduct that is intentionally, or at least consciously, motivated by race. That is, to label a conduct racially-motivated means that the intent attaches to an individual actor. Accordingly, the burden of finding a "single bad actor" has become a critical weakness as it is not always easy to spot a single responsible perpetrator (Cole, 1992:642; Foster, 1993:732). In most cases, several facilities, ranging from toxic factories to incinerators, may have been concentrated in one neighbourhood. This is why lawsuits are often filed against the permitting process instead of lodged against a particular facility. Moreover, if "the dynamics of a free market" causes racial inequities here, then the charge of environmental racism is likely to be dismissed by court since the bad "actor" is the market.

The second thing to be noticed is the limits of shifting the burden of proof. As Bullard (1999; 1995, 2000, 2001) has repeatedly argued, to prove intentional or purposeful discrimination in a court is next to impossible; therefore, the plaintiff should not be forced to shoulder the burden of proving a polluter's intent, i.e. shifting the burden of proof to polluters is essential. He proposes that at the very

12 *Id*. at 1150.

time companies apply for operating permits, they must "prove" not only that their operations are not harmful to human health, but also that these operations are not discriminatory and will not disproportionably impact minorities. The limitations on this argument are obvious. Just as the plaintiff cannot prove the defendant's intent, it is also next to impossible for a defendant to prove him- or herself innocent, or that their operations are non-discriminatory. On the contrary, suppose a defendant can and does prove that its operations have no disproportionate impacts on minorities at the time of site selection. If the unequal distribution is driven by market powers, rather than selection processes, over the long run there is little chance of keeping low-income and minority families from moving into the vicinity of these facilities. In short, shifting the burden of proof is no panacea for curing injustice.

Finally, the details of the analytic practices utilised in the court should never be overlooked. As we can see, the means of analysis seem be inconsistent in these cases. In *Bean*, several analytical units were used to construct defendant's racial intent: the entire city of Houston, the eastern and western halves of the city, specific tracts, and a target area. In *Bibb*, the court only accepted census tracts in its analysis. Finally, due to African American's extraordinary concentration around the target landfills, the court decided to use varying distances from the site as a device to inspect the distribution of a population in *R.I.S.E.* The question now arises: Is there a single most appropriate means for examining environmental racism or EJ?

In the following cases, I will concentrate on how the concept of EJ has been "evolving", if at all, as both as a conceptual matter and an empirical one. The inherent problems of EJ will be discussed in depth. It may be true, I argue, that the EJ terminology is evolving; the phenomenon itself remains contested.

3. Environmental equity: EJ in an income sense

Similar to the environmental racism debates, people continue to dispute or reinterpret the meaning of environmental equity. Yet, unlike environmental racism, which is mainly used in court or academia, federal agencies did once adopt it to guide their policies. This section explores why environmental equity was accepted at first but abandoned eventually by most activists and the US EPA.

3.1 Individualising EJ: Is inequity an individual choice?

When EJ concerns first came to federal decision-makers' attention in the early 1990s, the terminology that the EPA preferred was that of environmental equity rather than racism or justice. In January 1990, a conference entitled *Race and the Incidence of Environmental Hazards* was held at University of Michigan (Bunyan Bryant & Mohai, 1992a; 1992b; see also Reilly, 1992). During this conference, concerned activists and academics formed the so-called *Michigan Coalition*, and appealed to the EPA to address issues related to EJ. Before long, members of the Congressional Black Caucus and the Coalition met with EPA officials. In July 1990 the EPA Administrator created the Environmental Equity Workgroup to address the allegation that "racial minority and low-income populations bear a higher

environmental risk burden than the general population." (US EPA 2000) In 1992, this workgroup issued a report entitled *Environmental Equity: Reducing Risk for All Communities* (*Equity Report*) (US EPA, 1992a, 1992b).

This report raised the concerns about the access of racial and low-income groups to the policy-making process. It also considered the distributional issues of environmental problems in the light of race and income. Thus, for the first time, EJ's procedural and substantive dimensions were established. What is more, in this report the EPA made it very clear that the reasoning behind choosing *environmental equity*, instead of environmental justice or environmental racism, is scientific:

> [Environmental equity] most readily lends itself to scientific risk analysis. The distribution of environmental risks is often measurable and quantifiable. The Agency [EPA] can act on inequities based on scientific data. Evaluating the existence of injustices and racism is more difficult because they take into account socioeconomic factors in addition to the distribution of environmental benefits that are beyond the scope of this report. (US EPA, 1992a:10)

According to the *Equity Report*, air quality in Black and Hispanic communities did not meet federal standards and commercial hazardous waste facilities were more likely to be located in Black and low-income communities. Meanwhile, PCBs and dioxins are also more likely to accumulate in the bodies of racial minorities. The EPA concluded, that in most cases the lack of data and knowledge relating environmental health effects to race and income is a problem. Even so, it still documented that variations in exposure to lead related to income and racial factors. Their data showed a significantly higher percentage of Black children having high levels of lead in their blood (Reilly, 1992; US EPA, 1992a, 1992b, 1992d).

At first glance, it seems that this report has officially confirmed most of the concerns raised by the EJ movement, whereas quite the opposite is the case. The EPA implicitly adopted a "victim blaming" viewpoint, which considers *voluntary activities* the major source of causing disproportionate distributions (Foster, 1993:736). For instance, the EPA asserted that air pollution is "primarily an urban phenomenon, where emission densities tend to be the highest" (US EPA, 1992a:13); therefore, the hazards in African American's blood is simply because "[a] large proportion of racial minorities reside in metropolitan areas and may be systematically exposed to higher levels of certain air pollutants" (US EPA, 1992a:13).

Likewise, minorities' higher exposure to pollutants from toxic waste sites is simply caused by that fact that "minorities are more likely to live near a commercial or uncontrolled hazardous waste site" (US EPA, 1992b:7). For the same reason, one's exposure to contaminated fish may derive from their eating habits. Unsurprisingly, in the case of pesticides in Hispanic mothers' milk, the EPA indicated that:

> since racial and ethnic minorities comprise the majority of the documented

> and undocumented farm workforce, they may experience higher than average risk from agricultural chemicals (US EPA, 1992b:10).

The EPA's stance suggests that the distributional patterns of environmental problems in terms of race or income may have roots in one's choices of residence, job and diet:

> [I]t is becoming increasingly apparent that a person's activity pattern is the single most important determinant of environmental exposure for most pollutants. (US EPA, 1992b:7)

The blame, the EPA implies, falls to individual choices about where to live, work, and what to eat. Just as one cannot force smokers to quit, defining environmental equity as a matter of personal choice implies that there is nothing the EPA can do.

Unsurprisingly, this report's recommendations were primarily procedural; it's main focus is on how to involve more minorities or low-income groups in the decision-making system. Its most substantive suggestion is that more exposure data is needed, and its risk assessment process should be revised to incorporate the conception of environmental equity.

In sum, the term *environmental equity* was selected for its "scientific" nature. And, science/scientific assessment, the EPA suggests, is the best, if not the only way, to determine whether there are any population groups at disproportionately high risk. Finally, the *Equity Report* implied that environmental inequity may be caused primarily by personal choices.

3.2 What is wrong with the term of "equity"?

One of the *Equity Report*'s recommendations was to create mechanisms to tackle related problems; thus, in November 1992 the *Office of Environmental Equity* was established. Yet, is "equity", or as some called it the environmental (equity) paradigm (Foster 1993), a more desirable term for achieving EJ's goals?

3.2.1 From pattern seeking to harm assessing

To understand environmental paradigm, it is essential to identify the way that causation of harm is understood. The *Equity Report* conceptualises harm as the health impact resulting from any environmental contamination. In order to identify and measure harms, it is crucial to find a scientifically measurable link between exposure to hazardous materials and their impacts. The EPA does this by assessing the potential harms of a given substance and then setting a marginally acceptable level of safety to prevent physical environment and human health losses. In sum, the EPA seeks to prevent harms from happening by reducing them to the level where they have no detectable health effects (Foster 1993).

Unlike the remedial nature of the civil rights paradigm, under which a deprivation in relation to right is identified when a harm is caused by the "perpetrators", the environmental paradigm is, by nature, preventative. EJ advocators

need *not* identify a specific perpetrator or demonstrate one's discriminatory. Since health losses may not always be easily remediable, this preventative approach is considered much better suited to address injustice. Nevertheless, activists are still facing some great obstacles. Even if one could find sufficient evidence of racially disproportionate environmental hazard exposure, without scientific evidence to demonstrate a clear link between hazard and at least "potential" health harms, no harm is said to have occurred.

Risk, according to the EPA, is assessed in two steps. Firstly, a risk assessment is conducted. This process is almost exclusively dominated by a "scientific understanding of risk". Other factors, such as social or economic concerns, can only be considered once human health risk is identified, or the first step has been finished. It is in the risk management process, or the second step, where relevant equity issues should be evaluated. As the EPA can only "act on inequities based on scientific data" (US EPA, 1992a:10), a proper EPA response should only be considered during the second step, rather than in the beginning of a risk assessment.

Now, unless EJ advocates can "scientifically" demonstrate both the unequal distribution of hazardous facilities and actual health harms arising from the facilities, no remedial actions can be taken. Since most human effects of environmental harm are chronic and highly resource intensive, it is even harder and costlier to consider risks/harms than to simply demonstrate facilities' spatial distribution.

3.2.2 Does equity mean sending wastes to white communities?

Not only is the EPA's dependence on risk assessment problematic, the term *equity* is itself misleading as well. It distracts the aims of EJ from the prevention to the redistribution of pollution. In a redistribution model, so long as the toxic wastes are equally distributed, no matter how intensive the pollution is, equity can still be served.

As activists criticised both EPA's terminology and its dependence on risk assessment, the EPA soon accepted the more inclusive term *environmental justice* (Holifield, 2001:80). After President Clinton proclaimed Executive 12898 in 1994, environmental justice has been elevated to favoured term in the US federal government as a whole. Meanwhile, the name of Office of Environmental Equity was changed to *Office of Environmental Justice* (OEJ) in the same year (US EPA 2000).

In spite of the EPA's shift in terminology, some still use the earlier terms of environmental equity and environmental racism. For instance, Bullard considers environmental equity as a consequence of environmental racism, he stated:

> Blacks did not launch a frontal assault on environmental problems affecting their communities until these issues were couched in a civil rights context beginning in the early 1980s. They began to treat their struggle for environmental equity as a struggle against institutionalized racism and an extension of the quest for social justice. (Robert D. Bullard, 2000: 29).

Using environmental equity as a substitute for environmental racism is one efficient way to sidestep the almost insurmountable legal barrier of intent as the term of inequity often indicates *outcomes*, not the result of intent to harm (Rhodes, 2005:16-17). After losing the battle in *Bean*, it is not hard to know why Bullard made such an interpretation.

Other commentators (Holifield, 2001:80; Rhodes, 2005:16-17) criticise this term *environmental equity* for the reason that it lent itself to a scientific and redistributive understanding of EJ. Considering its political implications, they suggest that EJ research move toward a deeper understanding on structural injustice questions. Given that federal agencies and many activists now avoid the term, they further recommend that we drop it altogether.

4. Environmental justice: A better term than the others?

As environmental justice targets "any form of unequal distributed environmental hazard", it is arguably the most inclusive and most accepted term from both a social movement and a government agency perspective. Again, Bryant (1995) made a bid to capture its meaning:

> Environmental justice (EJ) [...] refers to those cultural norms and values, rules, regulations, behaviors, policies, and decisions to support sustainable communities where people can interact with confidence that the environment is safe, nurturing, and productive. Environmental justice is served when people can realize their highest potential [...] . EJ is supported by decent paying safe jobs; quality schools and recreation; decent housing and adequate health care; democratic decision-making and personal empowerment; and communities free of violence, drugs, and poverty. These are communities where both cultural and biological diversity are revered and highly reversed and where distributed justice prevails. (B. I. Bryant, 1995: 6)

From his definition, this expansive and more inclusive term includes provisions for both distributive and procedural justice. Distributive justice is concerned ensuring that no social group, no matter its socio-economic or racial character, suffers a disproportionate burden of negative environmental impact. Moreover, this substantive aspect of EJ emphasises the right to live in and enjoy a clean and healthful environment. On the other hand, procedural justice is concerned with ensuring that all communities have access to relevant information and claims that there should be a mechanism to allow locals to participate fully in decisions affecting their environment. (Agyeman & Evans, 2004; D. Schlosberg, 2007; US EPA, 2006). These two elements of EJ have appeared in official documents bearing on the subject.

4.1 No easy way to achieve EJ

Simply accepting EJ does not mean that it has paved an easier way for achieving it. The proposed Environmental Justice Act of 1992 (*EJ Act*)[13] provides an example of how hard it can be. Were it enacted, the EPA would have been required to identify the top 100 most-contaminated counties (other spatial units), known as Environmental High Impact Areas (EHIAs). This bill further called for study on evaluating "the nature and extent, if any, of acute and chronic impacts on human health" in EHIAs. If any significant impacts are found, a moratorium will be issued on the siting of any further facilities. Once the building of new facilities stopped, the accumulated quantities of hazardous substances will remain stable over the long run.

Obviously, *EJ Act*'s approach is neither remedial nor preventive in nature. It simply says that no more facilities in EHIAs; it however says nothing about where the next facility should be located or how to cure or compensate these areas. In addition, this approach evaluates only the burden of toxic chemicals and says nothing about non-noxious burdens, such as prisons or halfway houses.

This Act soon invites criticism. After being barred from siting new facilities in the top 100 EHIAs, new facilities may open in the 101-200 ranked areas. The second 100 EHIAs may have slightly less toxic exposure, but residents there may shoulder more non-noxious facilities, say prisons, than the top 100 EHIAs. The *EJ Act* approach offers no proper method to compare different burdens. More importantly, under this scheme some rich areas, say ranked 1,000th overall, may still shirk responsibility for sharing the burden of hosting any facilities (critique see: 1994a; Been, 1994b, 1994c).

By centring on a specific kind of burden, this Act addresses EJ in a scientifically limited way; this approach begs two fundamental questions: how to determine where the facilities are located, and how to measure the toxicity of a chemical and then compare the harm done with that due to others.

There is no straightforward means to demarcate the location of a facility. The *EJ Act* originally assumed a county-by-county designation, as demographic data is more available at jurisdiction levels. This Act however does not justify why political jurisdictions are more appropriate geographical units for measuring EHIAs. Given that facilities can easily be located on (even across) the border of several counties, a county-wide basis may not always reflect the actual impacted areas. Moreover, a jurisdiction's population may distribute unevenly. Without taking density into account, a less toxic but high population density facility may be left, while a highly polluted but low-density area may be categorised as EHIA.

Another benchmark case is the EO 12898, which marks the introduction of the term environmental justice into federal policy. This Order reinforced the 1964 Civil Rights Act Title (VI) that prohibits discriminatory practices and requires all federal agencies to begin to develop policies to promote a conception of EJ. The following discussion details the practical problem the EPA encountered in developing a programme to deliver EJ.

Despite the EO 12898 requiring federal agencies to ensure compliance with EJ requirements of the 1964 Civil Rights Act Title VI, the EPA had not produced any

13 H.R.2105, 103d Cong., 1st Sess. (1993).

program for violation of Title VI. In 1998, EPA's *Interim Guidance for Investigating Title VI Administrative Complaints Challenging Permits* (*Interim Guidance*) finally published. These guidelines established five steps in Disparate Impact Analysis. The first step is to identify the population affected by the permit, following which the racial and/or ethnic composition is to be determined. As there may be more than one facility in this area, the scope of facilities and total affected populations will be demarcated in the third step. After that, a disparate impact analysis will be conducted to compare the racial and ethnic characteristics within the impacted population. Finally, statistical analyses will be introduced to confirm whether the disparity is significant.

Again, the EPA faced a difficulty of determining whether its action constituted discrimination in violation of Title VI. To tackle this problem, the *Interim guidance* attempt to translate the legal term, discrimination, into a scientifically practical term, "disparate impact analysis". Surprisingly, the guidelines are particularly at a loss in conducting such as analysis.

A disparate impact analysis, as the EPA suggests, includes comparing the racial characteristics of the affected population to the non-affected. However, this guidance says nothing about how the communities to be compared will be identified, and how disparities between communities will be established. Eventually, even though such an analysis is conducted, it still does not provide any meaningful means to advance EJ. As before, this guidance also relies on a cumulative impact approach for determining environmental discrimination. This raises a new predicament: Should attention be focused on past, existing, or future injustice?

In the end, perhaps the actual analysis that these guidelines suggest are summarised in the fourth step:

> Since there is no one formula or analysis to be applied, the [EPA's Office of Civil Rights] may identify on a case-by-case basis other comparisons to determine disparate impact. (US EPA, 1998:10)

4.2 Beyond distributive justice?

In order to evaluate the distribution of environmental hazards, two approaches, forwarded on the basis of different conceptual frames, can be identified (cf. Been, 1994b; Helfand & Peyton, 1999:70). The "fair share approach" focuses on disparate exposure among various groups. Advocated by Bullard (1983), Gelobter (1992), and in part by the aforementioned *Interim Guidance*,[14] this frame accepts the hypothesis of "relative deprivation". The strictest interpretation of it implies that anyone, anywhere, should enjoy/share equivalent environmental quality/burdens; this approach therefore suggests that people assess their communities' environmental quality/burdens by comparison with their neighbours instead of against an absolute standard of living.

The other view, exemplifies by Bryant (1995) and to a certain extent by the

14 The reason why it is only partially adopted is because it is still uncertain how to choose a comparison community.

1992 *EJ Act15,* is "safe minimum standards approach". This approach does not ask for the same environmental quality in all communities, as does the fair share approach. Instead, EJ is judged by the fact whether or not the government achieves all kinds of environmental standards in all places. That is, meeting safety standards itself is achieving minimum EJ.

Without doubt, both approaches have political implications. The fair share approach attempts to describe differences by socioeconomic and ethnic factors first and then seek remedies; in contrast the minimum-standard approach asks authorities to identify key standards and areas not meeting those standards first and then requires an explanation for their shortcomings relative to socioeconomic or ethnic variables (Helfand & Peyton, 1999).

No matter which approach we prefer, it is evident that the meaning of EJ continues to be challenged. It may be true that comparing with other terms claiming to speak of EJ environmental justice is more comprehensive. However, adopting an all-embracing term will not automatically resolve the previously unsettled questions and therefore the broadly-defined EJ will inevitably bring its own methodological and philosophical baggage. For example, as in the *EJ Act*, most attention is given to the geographic location of toxic facilities, but other burdens are entirely left out of the equation. In turn, some EJ critics even find it difficult to call non-area/toxic-specific issues environmental "injustice" (Rhodes, 2005:27-28).

For me, instead of assuming that the claims of EJ refer to a universal, monolithic agenda, we should ask what environmental justice/equity/racism means in different contexts. The variety of definitions led me to conclude that "there is no such a thing as EJ", but there are many "EJs" (Poirier, 1994). To this end, an EJ scholar should provide the potential range and form of environmental-justice issues and, most importantly, to make the operative assumptions within each EJ frame explicit, rather than providing definitive categories (cf. Rhodes, 2005:29).

5. You have a dream; I have a problem: What does EJ really mean?

As I have revealed, people in different EJ periods have used environmental justice/equity/racism in varying ways, and the nature of terminology alone cannot account for the reasons why activists had abandoned one term and then chosen another. More than often, a term was abandoned for practical reasons. Environmental racism was the first term used to bring *justice* into the realm of environmentalism. Drawing on civil rights legacy, EJ advocates have successfully mobilised the impacted communities. However, as the courts have repeatedly rejected their EJ claims, activists soon noticed the limitation of environmental racism.

After EJ was written into the agendas of most environmental groups, activists no longer promoted the racial version of EJ, but adopted a new term environmental

15 This argument is made from the point of view that this act attempts to identify 100 EHIAs, or where the safe minimum stands have not been met. In other words, facilities should go into places meeting the minimum standards. However, having been discussed, we still do not know how to decide these standards or 100 EHIAs.

equity to integrate the poor into their schemes. Nevertheless, some still contend that environmental equity is inadmissible, due to its apparent emphasis on science and equitable redistribution of pollution. Being widely criticised, the EPA soon symbolically shifted its terminology to environmental justice.

While the term environmental justice was employed successfully as a rhetorical device to promote the EJ movement, academic debates continue. Seemingly, adopting an all-embracing term did not automatically settle the previous unsettled debates occurring in environmental racism and environmental equity phases.

The STS message that we can take with reference to EJ is that, we can understand a lot about the EJ movement by examining the social forces that shaped its development. Too often, as researchers have delved into the immediate demands of their respective domains of inquiry, these social factors have gone unnoticed. Without understanding this multifaceted nature and the social factors behind it, one is easily locked into the debate of which one definition is superior to all the others. This sense of superiority can easily cause a real bottleneck.

References

Agyeman, J., & Evans, B. (2004). 'Just sustainability': The emerging discourse of environmental justice in Britain? *Geographical Journal,* 170 (2), 155.

Balter, J. (2000). Environmental justice: Time for meaningful action. *Temple Environmental Law & Technology Journal,* 18 (2), 153-162.

Been, V. (1994a). Compensated siting proposals: Is it time to pay attention. *Fordham Urban Law Journal,* 21(3), 787-826.

Been, V. (1994b). Conceptions of fairness in proposals for facility siting. *Maryland Journal of Contemporary Legal Issues,* 5 (1), 13-24.

Been, V. (1994c). Siting of locally undesirable land uses: Directions for further research. *Maryland Journal of Contemporary Legal Issues,* 5 (1), 105-113.

Been, V., & Gupta, F. (1997). Coming to the nuisance or going to the barrios: A longitudinal analysis of environmental justice claims. *Ecology Law Quarterly,* 24(1), 1-56.

Bryant, B., & Mohai, P. (1992a). The Michigan Conference: A turning point. *EPA Journal, 18*(1), 9-10.

Bryant, B., & Mohai, P. (1992b). *Race and the incidence of environmental hazards: A time for discourse*. Boulder, CO: Westview Press Boulder.

Bryant, B. I. (1995). *Environmental justice: issues, policies, and solutions*. Washington, D.C.: Island Press.

Bullard, R. (1999). Leveling the playing field through environmental justice. *Vermont Law Review 23*(3), 453-478.

Bullard, R., & Wright, B. H. (1987). Environmentalism and the politics of equity: Emergent trends in the black community. *Mid-American Review of Sociology,* 12 (2), 21-38.

Bullard, R. D. (1983). Solid waste sites and the Black Houston community. *Sociological Inquiry,* 53 (Spring, 1983), 273-288.

Bullard, R. D. (1995). Decision Making. In L. Westra & P. S. Wenz (Eds.), *Faces of environmental racism: Confronting issues of global justice* (pp. 3-28). Lanham, Md.; London: Rowman & Littlefield.

Bullard, R. D. (2000). *Dumping in Dixie: race, class, and environmental quality* (3rd ed.). Boulder, Colo :: Westview Press.

Bullard, R. D. (2001). Environmental justice in the 21st Century: race still matters. *Phylon 49*(3/4), 151-171.

Bullard, R. D., Mohai, P., Saha, R., & Wright, B. (2007). *Toxic wastes and race at twenty 1987-2007: Grassroots struggles to dismantle environmental racism in the United States*. Cleveland, OH: United Church of Christ Justice.

Capek, S. M. (1993). The "Environmental Justice" Frame: A Conceptual Discussion and an Application. *Social Problems,* 40 (1, Special Issue on Environmental Justice), 5-24.

Chavis, B. (1994). Preface. In R. D. Bullard (Ed.), *Unequal protection: environmental justice and communities of color* (pp. xi-xii). San Francisco: Sierra Club Books.

Cole, L. W. (1992). Empowerment as the key to environmental protection: The need for environmental poverty law. *Ecology Law Quarterly* 19 (4), 619-683.

Collin, R. W. (1992). Environmental equity: A law and planning approach to environmental racism. *Virginia Environmental Law Journal,* 11 (4), 485-546.

Dobson, A. (1998). *Justice and the environment: Conceptions of environmental sustainability and theories of distributive justice*. Oxford: Oxford University Press.

Foreman Jr, C. H. (1998). *The promise and peril of environmental justice*. Washington, D.C.: Brookings Institution.

Foster, S. (1993). Race(ial) matters: The quest for environmental justice. *Ecology Law Quarterly,* 20 (4), 721-753.

Gelobter, M. (1992). Toward a model of "Environmental Discrimination". In B. Bryant & P. Mohai (Eds.), *Race and the incidence of environmental hazards* (pp. 64-81). Boulder, CO: Westview Press.

Gleeson, B., & Low, N. (2003). Environmental justice. In J. Agnew, K. Mitchell & G. Toal (Eds.), *A companion to political geography* (pp. 455-469). Oxford: Blackwell.

Harvey, D. (1996). *Justice, nature and the geography of difference*. Oxford: Blackwell.

Harvey, D. (1999). The environment of justice. In F. Fischer & M. Hajer (Eds.), *Living with Nature* (pp. 153-186). Oxford: Oxford University Press.

Helfand, G. E., & Peyton, L. J. (1999). A conceptual model of environmental justice. *Social Science Quarterly, 80*(1), 68-83.

Hoban, T. M., & Brooks, R. O. (1996). Environmental ideals: Equitable distribution of rights and liabilities. In T. M. Hoban & R. O. Brooks (Eds.), *Green justice: the environment and the courts* (2nd ed., pp. 157-168). Boulder: WestviewPress.

Holifield, R. (2001). Defining environmental justice and environmental racism. *Urban Geography,* 22 (1), 78-90.

Lavelle, M., & Coyle, M. (1992). Unequal protection: The racial divide in environmental law. *National Law Journal,* 15 (3), S1-S12.

Lyle, J. M. (2000). Reactions to EPA's Interim Guidance: The growing battle for control over environmental justice decision making. *Indiana Law Journal,* 75 (2), 687-708.

MacGregor, D. G., Slovic, P., & Malmfors, T. (1999). How exposed Is exposed enough?

Lay inferences about chemical exposure *Risk Analysis,* 19(4), 649-659.
Maher, T. (1998). Environmental Oppression: Who Is Targeted for Toxic Exposure? *Journal of Black Studies, 28*(3), 357-367.
Pellow, D. N. (2000). Environmental inequality formation: Toward a theory of environmental injustice. *American Behavioral Scientist* (4), 581-601.
Poirier, M. R. (1994). Environmental Justice / Racism / Equity: Can we talk? *West Virginia Law Review,* 96 (1083), 1083-1107.
Pulido, L. (1996). A critical review of the methodology of environmental racism research. *Antipode,* 28 (2), 142-159.
Reilly, W. K. (1992). Environmental equity: EPA's position *EPA Journal,* 18 (1), 18-22.
Rhodes, E. L. (2005). *Environmental justice in America: A new paradigm*. Bloomington, IN: Indiana University Press.
Ringquist, E. J. (2006). Environmental justice: Normative concerns, empirical evidence, and government action. In N. J. Vig & M. E. Kraft (Eds.), *Environmental policy: New directions for the twenty-first century* (6th ed., pp. 249-273). Washington, D.C.: CQ Press.
Schelly, D., & Stretesky, P. B. (2009). An analysis of the "Path of Least Resistance" argument in three environmental justice success cases. *Society & Natural Resources,* 22(4), 369-380.
Schlosberg, D. (1999). *Environmental justice and the new pluralism: The challenge of difference for environmentalism*. Oxford: Oxford University Press.
Schlosberg, D. (2007). *Defining environmental justice: Theories, movements, and nature*. Oxford: Oxford University Press.
Taylor, D. E. (2000). The rise of the environmental justice paradigm: Injustice framing and the social construction of environmental discourses. *American Behavioral Scientist* (4), 508-580.
Tesh, S. N., & Williams, B. A. (1996). Identity politics, disinterested politics, and environmental justice. *Polity,* 28 (3), 285-305.
The First National People of Color Environmental Leadership Summit. (1991). *Principles of Environmental Justice.* Paper presented at the the First National People of Color Environmental Leadership Summit, , Washinfton, DC.
Tsao, N. (1992). Ameliorating environmental racism: A citizens' guide to combating the discriminatory siting of toxic waste dumps. *New York University Law Review,* 67(2), 366-418.
US EPA. (1992a). Environmental equity: Reducing risk for all communities. Volume1: Workgroup report to the administrator, from http://www.p2pays.org/ref/32/31476.pdf
US EPA. (1992b). Environmental equity: Reducing risk for all communities. Volume2: Supporting document, from http://www.p2pays.org/ref/32/31476.pdf
US EPA. (1992c). History of Environmental Justice: Basic Information Retrieved 09/29, 2006, from http://www.epa.gov/compliance/basics/ejbackground.html
US EPA. (1992d). Release of environmental equity report Retrieved May 20, 2009, from http://www.epa.gov/history/topics/justice/01.htm
US EPA. (1998). *Interim Guidance for Investigating Title VI Administrative Complaints Challenging Permits*.
US EPA. (2006). Environmental for all (確保人人享有環境公正). from http://www.epa.

gov/chinese/pdfs/EJ%20Brochure_CHI.pdf

US EPA. (2009d). Environmental justice: Frequently asked questions Retrieved Nov 20, 2009, from http://www.epa.gov/compliance/resources/faqs/ej/index.html

US EPA, U. E. P. A. (2000). What's New: Environmental Justice Map Protocol Available from www.epa.gov/region4/ead/WhatsNew/WhatsNew.html

Zimmerman, R. (1993). Social Equity and Environmental Risk1. *Risk Analysis,* 13 (6), 649-666.

Epicureanism about the Badness of Death and Experientialism about Goodness

Fumitake YOSHIZAWA

1. Introduction

Is death bad for the one who dies? The position that answers this question in the negative is called—in the broadest sense of the word—*Epicureanism* about the badness of death.[1] The purpose of this paper is, first, to try to clarify a certain type of criticism of Epicureanism, namely, that Epicureanism (the denial of the badness of death) seems to be incompatible with the affirmation of the goodness of life (*experientialism about goodness*). Second, I try to make a partial defense of Epicureanism by responding to the criticism.

The paper proceeds as follows: To begin with, I clarify some possible misunderstandings regarding Epicureanism. Next, I formulate several versions of what I refer to as the *Argument from Experientialism about Goodness* (AEG), which attempts to point out the difficulty of Epicureanism based on experientialism about goodness. I claim that all the versions are unsound; each version is either unfair to Epicureanism or can be responded to by endorsing an available proposal that is based on the criticism of *comparativism* about prudential value. In addition, I attempt to make a proposal that explains how we (typically) do *not* choose rationally to die, without presuming the badness of death. Finally, I try to dispel some possible worries about my proposal.

2. Preliminary

Let us begin by clearing up two possible misunderstandings that make Epicureanism look "absurd." First, in order to understand Epicureanism properly, we have to distinguish *dying* from *being dead. Dying* is the last process of *life,* whereas *being dead* is the duration *after life has ended*.[2] Epicureanism is *not* the view that asserts that *dying* is free from any badness. According to Epicureanism, for example, pains

1 The purpose of this paper is not to provide an accurate interpretation of Epicurus' original text but to examine "Epicureanism" in the broad sense used in contemporary analytic philosophy.

2 This is a very rough characterization for the sake of simplicity. In particular, it is more precise to define "dying" with a phrase like "wherein certain causes operate to bring about one's death" (Li 2002, 13). Dying typically involves certain biological signs. It might be also possible to define the word "death" as, so to speak, the boundary between *dying* and *being dead* (cf. Rosenbaum 1986, 217-8). There is no need to assume, however, that there is an intervening *stage*, not just a *boundary*, between dying and being dead. Indeed, there is even a problem in making such an assumption (see Li 2002, 13-6). In this paper, I use "death" to mean the state of affairs of being dead.

and sufferings involving dying with a disease *are* bad. In the most usual situations about which we would say, "death is bad," the Epicurean is also able to admit that the course to (or process of) death is bad. According to Epicureanism, however, that which is bad in such a situation is not death *per se,* namely, *being dead,* but the pains involved in the process, namely, *dying*.

On another possible misunderstanding: Epicureanism is *not* the view that insists that a short-term painful (thus, *bad*) experience like a stomachache is worse than painless (thus, *not bad*) death. This kind of misunderstanding about Epicureanism seems to come from mistaking the relevant objects of the comparison. Suppose that there are two possibilities: in one possible state of affairs, you will experience an intense stomachache for 10 minutes from the time *t*, and in the other possible state, you will die at the time *t*. According to Epicureanism, the experience of the stomachache is worse than the state of "being dead" just for those 10 minutes when the stomachache continues. This is because death has no value. Death is *not,* however, like a state of no consciousness that continues only for 10 minutes (what we might think of as "10 minutes of death" is not death at all, but 10 minutes without consciousness). In such a situation, what should be compared with the state of being dead for a person after time *t* is the counterfactual whole remaining course of life after *t* that contained the experience of the 10 minutes of intense stomachache. The duration of being dead after time *t* has no value and then, in most cases, is *worse* than the counterfactual course that we would continue to live (even with some periods of suffering). The Epicurean can also admit to this evaluation. Nevertheless, the Epicurean denies the evaluation of the *badness* of death, which is made on the ground of the fact of being merely *less good* (*worse*) than the alternative. I return to discuss this point in detail later.[3]

3. The argument from experientialism about goodness (AEG)

Experientialism about badness is essential for the Epicurean argument against the badness of death. Experientialism is an axiological theory about the welfare (or well-being) of persons. According to the theory, only bad experiences, such as pain or suffering — or, more precisely, the state of affairs of one's experiencing the bad sensations[4] — are the bearer of badness for a person.[5] Experientialism about badness

3 This paper intends to present a partial defense of Epicureanism. There are indeed several problems for Epicureanism that I do not deal with in this paper. It is often pointed out that if death is not bad, one cannot explain the wrongness of killing others. There is, however, a persuasive response (Burley 2010). In any case, the problem that I take up in this paper is not this, but the (alleged) problem of an intrinsic inconsistency in Epicureanism.

4 In this paper, I talk of the value of *states of affairs*, not the value of *events*. If I understand it correctly, the evaluation of an event is based on the value of the states of affairs that are the outcomes of the event, so there is little importance in whether we should talk of states of affairs or events. In this paper, I take states of affairs as the bearer of value for a person. Thus, Epicureanism is the view according to which the state of affairs of one's being dead is not bad for the one who dies.

5 The name "experientialism" is from Soll (1998) and Schumacher (2010, chapter 9).

seems to assure that the state of affairs of being dead is not bad for the one who dies. Because death is the end of one's life, and so is the end of one's experiences, the state of being dead does not involve any bad experiences for the one who dies.[6]

I call in this paper the following type of argument the AEG: If you admit experientialism about badness, it is natural to admit the counterpart claim about goodness as well. Indeed, the best-known (and tenable) type of *hedonism* makes claims about both goodness (i.e., pleasure) and badness (i.e., pain). According to experientialism about goodness, parallel to experientialism about badness, good things for a subject consist in the subject's good experiences. Accordingly, the subject's *being alive* is necessary to realize any goodness for the subject. Does this just mean, after all, that *not being alive,* or in other words, *being dead,* is bad for the one who dies? Are the contentions that being alive is good and that being dead is *not bad* compatible? In order to defend Epicureanism, should any goodness be given up? This is a very rough characterization of an AEG.[7]

It may be helpful to provide some supplementary remarks. For the above type of argument, if the reasoning is valid, the argument that *there is a case* where being alive is good (that is, *sometimes* being alive is good) is enough to stand as a criticism of Epicureanism. Since the Epicurean says that deaths are *always* not bad, if one can conclude from the AEG that there is a case where being dead is bad from an existing case of a good life, then Epicureanism turns out to be false. Taking into account this point, the schema of the AEG can be restated as follows:

The AEG

1. In cases where there are good experiences, being alive is good. [From experientialism about goodness]
2. Hence, there is a case where being alive is good. [From 1]
3. [...]
4. Hence, there is a case where being dead is bad. [From 2 and 3]
5. Therefore, the claim of Epicureanism that "deaths are always not bad" is false.

There are several ways of filling up the blank [...], that is, the alleged premise 3 from which the sub-conclusion 4 follows with the premise 2. In what follows, I examine several possible proposals of filling the blank and then show that the Epicurean can respond to each version of the AEG.

However, Soll's use of the term carries with it implications about motivation.

6 For the proposal to defend Epicureanism by showing the plausibility of a version of experientialism itself, see Rosenbaum (1986).

7 For this line of objection, which points out the apparent incompatibility of Epicureanism and the goodness of pleasure (or goodness of desire-fulfillment), see Miller (1976, 171), Warren (2004, 199-212), and Olson (2012, esp. section 6).

4. Meaning

One might think that death being bad is a direct entailment of life being good. The blank [...] might be filled with the following assumption:

(**Meaning**)
3. "Being alive is *good*" *means* "being dead is *bad.*"

Therefore, there is a case where being dead is bad and so Epicureanism is wrong.

If there will be a good experience in some near future time, the state of being alive at that time is good for you. Because "being alive" means "being not dead," being *not* dead is *good* (at that time). Furthermore, that being *not* dead is *good* means that being dead is *bad*. One might claim that the conceptual truth of goodness and badness entails this. If it is true, it is the most crucial criticism of Epicureanism. Nevertheless, it is not true. The first half is fine. For an already-existing person, being alive is equivalent to being not dead. The second half, however, is not generally true. For example, while staying home tomorrow is good for you, being out (*not* staying home) also might be good.[8]

How about the following?

(**Meaning'**)
3a. Being alive is a necessary condition for any goodness.
[From experientialism about goodness]
3b. Hence, being not alive (being dead) is not good.
3c. "Being dead is *not good*" means "being dead is *bad*."

Therefore, (there is a case where) being dead is bad.

The Epicurean can admit 3a and 3b as implications of experientialism about goodness, but in order to make 3c correct, another assumption is required, that is, the assumption that each state of affairs is either good or bad for a particular subject.[9] Making such an assumption is unfair to the Epicurean, however. The Epicurean can just deny this assumption. Indeed, usually the Epicurean asserts the view that the state of being dead is neither good nor bad (in itself) for the one who dies. There is no constraint on the Epicurean to admit 3c, and she can claim that "being not good" does not mean "being bad" without any problem. Therefore, (**Meaning'**) also does not succeed (and furthermore, the view that the state of being dead with no experiences has no value [in itself] is necessary to claim experientialism about both badness and goodness consistently—needless to say, it should be a premise of any AEG).

8 "*P* is *good* for a subject if and only if *not P* is *bad* for the subject" is *not* generally true [where *P* stands for a given proposition].

9 On this assumption, "*P* is *not good* for a subject if and only if *P* is *bad* for the subject" is true.

5. Comparison

One might claim that the reasoning that fills the blank [...] is the following:

(**Comparison**)
3a'. Being dead has no value in itself.
[From experientialism about both badness and goodness]
3b'. When being alive is good, in comparison between being alive and being dead, being alive is *better* and being dead is *worse*.
3c'. Hence, in such a case, being dead is *extrinsically* bad.

In other words, there is a case where being dead is "bad," although it is not intrinsically bad. Here, the issue is the plausibility of *comparativism*. The most popular view about the badness of death, the so-called *deprivation account,*[10] is indeed comparativism about *extrinsic value* for a person. Thus, the AEG turns out to be an argument that attempts to show the difficulty of Epicureanism through the support of comparativism.

First, it is possible to put up the following counterargument against comparativism as a reply: Since death is an experiential blank, it is neither a good experience nor a bad experience. It should, however, be distinguished from an *experience* that is neither good nor bad. Because the one who died no longer exists, there seems to be a serious difficulty in attributing any value (even zero value) to the state of her being dead—or, the value of death must be left undefined. If no value can be attributed to the state of being dead, the comparison with other states cannot be established. To put this problem in a more formal way, there is a difficulty in ascribing *the property of having a well-being level* to the one who no longer exists.[11]

10 For the advocates of the deprivation account, see, for example, Nagel (1979), Feldman (1991), Feit (2002), Luper (2009), and Bradley (2009). There are at least two famous problems with this view that I do not take up in this paper: (1) the "problem of preemption"—the deprivation account seems not to explain the badness of death in cases where, if someone had not died by the cause by which he actually died, he would have died by another cause (McMahan 1988, 45); and (2) the "symmetry argument"—the deprivation account seems to admit the (counterintuitive) badness of pre-vital deprivation (for an attempt to solve the problem, see Kaufman 1999).

11 This problem is known as the "*no-subject* problem" or "*missing subject* problem" (or something similar). More properly, an abstract principle of the *existential requirement for badness* works here; the principle says that for something to be bad for a subject at a time, it is necessary for the subject to exist at that time. The *termination thesis*, which claims that when a person dies the person ceases to exist (cf. Feldman 2000, esp. 100) also works here. I take the termination thesis as a given. Experientialism about badness may be one of the simplest theories of welfare with the existential requirement. For criticism of the ascription of any level of welfare to a nonexistent dead person, see Silverstein (1980, 410-3), Silverstein (2010, esp. 290), and Luper (2007, esp. 247). For an attempt at a proposed solution by metaphysical frameworks, see Yourgrau (1987) and Ruben (1988). However, it is necessary to argue whether the ascription of *values* to the dead is possible or not, in addition

Comparativism should reply to this counterargument.

Even if the Epicurean admits the attribution of (zero) value to the state of nonexistence and thus admits the comparison of values, she can again criticize comparativism as a reply to the above form of AEG. Aaron Smuts makes such an argument against comparativism. The point of Smuts's argument is that the comparativist "conflates things that are merely less good with those that are bad" (Smuts 2012, 198). According to the comparativist, "deprivations" of goodness—that is, goodness's having *possibly* occurred but actually having *not* occurred—are *bad*. That is, in this view, the value of a state of affairs is determined on the basis of the comparison between the value of the actual state and the value of the counterfactual state—*less good* (or more bad) states *are* extrinsically *bad,* and *less bad* (or more good) states *are* extrinsically *good*. Smuts points out that admitting the extrinsic (comparative) value is problematic and advocates a *non*-comparativistic view—he claims that only things that are intrinsically bad and those that lead to intrinsic badness are bad.

One can find Smuts's criticism convincing by considering the following example.[12] Suppose that there are two chocolate chip cookies and that the right one contains one more chocolate chip than the left one. By choosing the left one, you eat fewer chocolate chips. According to comparativism, the situation in which you eat fewer chocolate chips is *bad.* Contrary to comparativism, however, the situation is obviously *good*; *not bad*, just *less good*. The cookie that you eat is delicious enough, and it brings a good experience to you. With respect to badness, according to comparativism, the merely *less bad* situation is *good.* For example, when a severe pain and a less severe pain are compared, according to this view, the less severe pain (it is obviously a *bad* experience) is *good*. Again, contrary to comparativism, the world is not "such a nice place that the best option was always a good one" (Smuts 2012, 209).

Consider another example: Thanks to a rich person's whim, you suddenly encounter the chance to get one of two briefcases with wads of cash. The right one that you get as a result contains $100,000, and you become extremely happy. You never know that the left one that you do not choose contains $100,000,000. In such a case, do you consider that a *bad* thing has happened to you? According to comparativism, choosing the right one is indeed *bad* for you. Nevertheless, the state of affairs of getting $100,000, which your choice leads to, is just *less good* than the other. It is never *bad*.[13]

It is worth noting that almost the same argument can be applied to the comparativistic view about *harm* to others—the view that identifies *harming* others with *making* others *worse off*.[14] Moreover, correspondingly, comparativism seems

to defending the general metaphysical frameworks of ascription of properties (even if the termination thesis is not true as Feldman himself claims, the plausibility of ascription of welfare to the dead—for example the existing dead *body*—should be explained). A related problem arises about the morality of procreation (see note 14 below).

12 I simplify Smuts's original example (Smuts 2012, 208).

13 I made a number of changes to Smuts's original example (Smuts 2012, 205-6).

14 For example, Feinberg (1984, 33-4) and Parfit (1984, 69). Luper (2009) proposes a uniform

to produce an intuitively wrong result about what we morally ought not to do to others, and thus about what actions are *morally wrong*. That is to say, according to comparativism, doing a less good thing *is* harm and therefore wrongdoing.[15,16]

Of course, we should be careful about the terminology. The comparativist makes a claim about *extrinsic value* for a person, not just about *intrinsic value* for a person. The comparativist might say that the above (alleged) problem of comparativism seems to arise from conflating a claim about extrinsic value with a claim about intrinsic value. The point is this, however: the comparativist does not just call a comparative value "extrinsic value" but also asserts that the comparative value is an important thing for a person. She says the comparison additionally determines significant things for us (that is, well-being). The Epicurean can accept the comparison itself but also claims it is *in*significant. The disagreement between these two views is thus substantial. Smuts's argument shows that comparativism results in some counterintuitive consequences, and there seems to be no advantageous reason to support the view over the *non*-comparativistic view, which the Epicurean stands for.

As seen above, (**Comparison**) is problematic and the Epicurean need not admit this version of AEG.

---

analysis of the badness of death by means of the concept of harm. One might claim that there are both intrinsic and extrinsic harms, but the problem turns out to be whether the comparative extrinsic harms have substantial roles or significance for us. For a defense of a *non*-comparative account of harm in the context of the problems of the morality of procreation, in particular the "paradox of future individuals" (Kavka 1982) otherwise well known as the "non-identity problem" (Parfit 1984, 359), see, for example, Harman (2004). It is also worth noting that, as I understand it, one of the crucial points of the non-identity problem is essentially the same as the point of no-subject problem (see note 11, above). The common point is that there seem to be no relevant comparison of values for a person. The non-identity problem is that it seems impossible to harm (or benefit) offspring by procreating them. According to comparativism, to harm (or benefit) someone, it is necessary to make them worse off (or better off), but in the situation of procreation the comparison cannot be established because if a (possible) child has not been brought into existence, then the child has no level (even zero level) of well-being at all.

15 Conversely, according to comparativism, an action preventing other persons from suffering worse consequences makes the action *harmless*. For example, according to comparativism, a man who is stopped from boarding a flight because of his race *benefits* from racial discrimination if that flight crashes (Woodward 1986, 810-1).

16 One might think that all these counterexamples are avoided by claiming that worseness (betterness) is just a *necessary condition* of extrinsic badness (goodness), harm (benefit), or wrongness (rightness)—perhaps, by additionally claiming that actual intrinsic badness with no alternative badness is a *sufficient condition* of extrinsic badness. This position fails to explain what the badness of death consists in, however, because death is an experiential blank and there are no grounds for the ascription of values to death except for being comparatively less good.

6. Reason

One might think that (**Comparison**) still has some appeal or is even indispensable for us. Consider, however, the following imagined situation: There are two one-and-a-half-month-old kittens. Kitten A, with a weight of 2 pounds, is heavier than Kitten B, which weighs 1 pound. Kitten A is, so to speak, *extrinsically heavy.* Well, so what? Kitten A may be healthy and Kitten B may not, but Kitten A is healthy by virtue of weighing 2 pounds, *not* by virtue of being heavier than Kitten B. In this kind of "comparison," there seems no additional importance. Again, what is extrinsic value after all?

Perhaps, one might think that (**Comparison**) appears to be important for us because we indeed act according to comparative evaluations in some situations.[17] Furthermore, one might think that if one's death is sometimes rationally *avoided,* and if it is possible that one is able rationally *not* to die, the following should be true as the premise coming after (**Comparison**):

(**Reason**)

3d'. Therefore, in such case, there is a *reason in favor of* the state of being alive, and there is a *reason against* the state of being dead.

Therefore, the aim of AEG is now not to show the incompatibility between Epicureanism and the experientialism about goodness, but to show the unacceptable consequences from the combination of these views. If (**Reason**) is not true, however, is death always rationally to be brought about? Is (**Reason**) the only way to keep off such an all too unacceptable consequence? The answer is no. Indeed, it might be true that (**Reason**) is required for (literal) justification of *avoiding* death, but the all too unacceptable consequence can be resisted by the following, which does not assume (**Comparison**):

(**Reason'**)

3a'. Being dead has no value in itself. [From experientialism about both badness and goodness]

3b". Therefore, there is a case where there is a reason in favor of the state of being alive and there is *no reason* in favor of the state of being dead.

What is important here is this: the mere fact that within the two options, one of them provides a prudential reason in favor of a state of affairs and the other does *not*, does *not* provide a reason against or a reason to avoid the other state of affairs. We just

17 Smuts argues on the same point that we indeed make comparative evaluations and choose an available option accordingly. He concludes that our comparative evaluations are not relevant because merely less good outcomes that result from our choices are not, thereby, bad (Smuts 2012, 207-9). I agree with his conclusion itself, but I think there remains a significant feature of our choices that our comparative evaluations seem to track. In this section, I argue about this feature, namely, the relation that our choices have with values and reasons.

have *no reason* to bring about or to avoid that other state, whereas the first state of affairs just gives us sufficient prudential reason to bring it about. The results of the action that is pursued (with success) according to prudential reasons may indeed fit the comparative evaluation. A rational person, however, can just pursue continuing to live without having reason to avoid death. Again, the latter state of affairs with no reason in favor of it is *not bad* but merely *less good.*

This line of thought about reasons is important when the first criticism against comparativism mentioned in section 5 is admitted; that is, that death has no value and the comparison between the values of death and living alternatives cannot be made. According to this line of thought about reasons, even when the comparison of *values* is impossible, a sufficient reason for action is given to the agent in these situations and the agent can make her decision. Thus, for example, if death provides you with no reason to pursue it while the alternative state of affairs, that you continue to live, provides you with a reason to pursue it, then you can rationally decide to continue to live.[18]

Still, the Epicurean should perhaps admit that death has a *negative* aspect in the following sense: In many ordinary cases, the state of being dead provides *no reason* to bring it about and is not chosen because there are many rival alternative reasons. If there is a rival reason, all other things being equal, *not* acting according to the sufficient reason is (prudentially speaking) *irrational.* In such a case, to die is an *irrational* choice. This, again, should be distinguished from the position that death is bad or there is a reason *to avoid* death.[19] It might be that the Epicurean should refrain from saying, "death is nothing to us," and it is possible to say that there is a negative aspect to death, but this never means that death is bad in the sense that people ordinarily mean when referring to the badness of death.

7. Conclusion

I conclude that all versions of the AEG examined above do not succeed. The denial of the badness of death and experientialism about goodness are compatible, so the Epicurean does not have to give up goodness.

In closing this paper, let me take up two possible worries about my proposal. The first one concerns the claim that death cannot be good. According to my proposal, this means that there is no reason in favor of death. Someone might say, however, that there seems to be a case that death *does* provide a reason to choose it. For example, one might say, there is a reason to choose euthanasia over continuing

18 For an argument defending the possibility of practical reasoning that is not necessarily based on comparison, see Chan (2010). According to Chan, there is no requirement to believe that "any choice made with the capacity to reason requires the comparing of alternatives" and that to think that rational choice requires maximizing a choice value is a narrow view of rationality (Chan 2010, esp. 150-1).

19 As Smuts points out, the emotion of *fearing* seems to be appropriate not on the basis of the mere worseness of the situation but on the basis of the badness of the situation (Smuts 2012, 210-1).

unrecoverable intense pain. A response to this concern can be given as follows: Even in such a severe situation, euthanasia is in fact *not* a good choice; there is always a reluctance to choose it. The consequence of my proposal that there is no reason in favor of euthanasia tends to accord with our understanding of it. By contrast, according to comparativism, there *is* a reason in favor of euthanasia, and on this point, there is little intuitive appeal to admit comparativism.[20]

The second possible worry concerns the claim that death is not bad. This means that there is no reason to avoid death. First, the process of dying typically involves bad experiences, such as pains and sufferings (as I said in section 2, dying is typically bad). This explains the fact that we are *typically* inclined to avoid death.[21] Or someone might say the following: in situations when short-term bad experiences are expected, if there is a reason to avoid bad experiences and there is no reason to avoid death, then, does this mean that choosing death would be rational? As I said in section 2, however, we should be careful about the appropriate alternative option that should be compared with death. When the comparison is made appropriately, most ordinary cases show that there is a reason to opt for continuing to live (otherwise, euthanasia would have to be deliberately considered). The position of this paper does not entail the unacceptable consequence such that in everyday life death should be rationally chosen.

Acknowledgements

I would like to thank especially the following people for helpful discussion and comments: Jack Lee, Fan-Ching Leung, Kazunobu Narita, Ikuro Suzuki, and Taku Tanikawa. I also thank the audience at the Seventh International Conference on Applied Ethics. This work was supported by Grant-in-Aid for JSPS Fellows Number 11J03001.

References

Bradley, B. (2009) *Well-Being and Death*, Oxford: Clarendon Press.

Burley, M. (2010) "Epicurus, Death, and the Wrongness of Killing," *Inquiry: An Interdisciplinary Journal of Philosophy*, 53 (1), 68-86.

Chan, D. K. (2010) "Reasoning without Comparing," *American Philosophical Quarterly*, 47 (2), 149-60.

20 Remember Smuts's criticism: comparativism makes things that are just "less bad" good (Smuts 2012, 216).

21 I think there is, in fact, a disagreement over the matter of painless death, as Nagel says in the beginning of his remarkable work "Death" (Nagel 1979, 1). Although some of us indeed have the intuition that death is terribly bad (an anti-Epicurean way of thinking), others have pre-theoretical understandings such that "death has no value whatever, positive or negative" (a pro-Epicurean way of thinking). Moreover, people belonging the latter camp have no objection to death *per se*. Sumner also agrees that he belongs to the camp who denies that the dead can be either benefited or harmed (Sumner 1996, 127).

Feinberg, J. (1984) *Harm to Others*, Oxford: Oxford University Press.
Feit, N. (2002) "The Time of Death's Misfortune," *Noûs*, 36 (3), 359-83.
Feldman, F. (1991) "Some Puzzles About the Evil of Death," *Philosophical Review*, 100 (2), 205-27.
Feldman, F. (2000) "The Termination Thesis," *Midwest Studies in Philosophy*, 24, 98-115.
Harman, E. (2004) "Can We Harm and Benefit in Creating?" *Philosophical Perspectives*, 18 (1), 89-113.
Kaufman, F. (1999) "Pre-vital and Post-Mortem Non-Existence," *American Philosophical Quarterly*, 36 (1), 1-19.
Kavka, G. S. (1982) "The Paradox of Future Individuals," *Philosophy & Public Affairs*, 11(2), 93-112.
Li, J. (2002) *Can Death Be a Harm to the Person Who Dies?*, Dordrecht: Kluwer Academic Publishers.
Luper, S. (2007) "Mortal Harm," *Philosophical Quarterly*, 57 (227), 239-51.
Luper, S. (2009) *The Philosophy of Death*, Cambridge: Cambridge University Press.
McMahan, J. (1988) "Death and the Value of Life," *Ethics*, 99 (1), 32-61.
Miller Jr., F. D. (1976) "Epicurus on the Art of Dying," *Southern Journal of Philosophy*, 14 (2), 169-77.
Nagel, T. (1979) "Death," in *Mortal Questions*, Cambridge: Cambridge University Press, 1-10.
Olson, E. T. (2012) "The Epicurean View of Death," citing advance online publication: *Journal of Ethics*, doi: 10.1007/s10892-012-9140-8.
Parfit, D. (1984) *Reasons and Persons*, Oxford: Clarendon Press.
Rosenbaum, S. E. (1986) "How to Be Dead and Not Care: A Defense of Epicurus," *American Philosophical Quarterly*, 23 (2), 217-25.
Ruben, D-H. (1988) "A Puzzle about Posthumous Predication," *Philosophical Review*, 97 (2), 211-36.
Schumacher, B. N. (2010) *Death and Mortality in Contemporary Philosophy*, Cambridge: Cambridge University Press.
Silverstein, H. S. (1980) "The Evil of Death," *Journal of Philosophy*, 77 (7), 401-24.
Silverstein, H. S. (2010) "The Time of the Evil of Death," in *Time and Identity*, J. K. Campbell, M. O'Rourke & H. S. Silverstein (eds.), Cambridge, MA: MIT Press.
Smuts, A. (2012) "Less Good but not Bad: In Defense of Epicureanism about Death," *Pacific Philosophical Quarterly*, 93 (2), 197-227.
Soll, I. (1998) "On the Purported Insignificance of Death: Whistling before the Dark?" in *Death and Philosophy*, J. Malpas & R. C. Solomon (eds.), London and New York: Routledge.
Sumner, L. W. (1996) *Welfare, Happiness, and Ethics*, Oxford: Clarendon Press.
Warren, J. (2004) *Facing Death: Epicurus and His Critics*, Oxford: Clarendon Press.
Woodward, J. (1986) "The Non-Identity Problem," *Ethics*, 96 (4), 804-31.
Yourgrau, P. (1987) "The Dead," *Journal of Philosophy*, 84 (2), 84-101.

Liberty and Freedom: The Relationship of Enablement

Michael YUDANIN

1. Introduction: Freedom and liberty

If freedom is a capacity of embodied reason, something we have by virtue of being humans, how should it impact society? The liberal view that subscribes to the notion of universal human rights would claim that political liberty is a societal accommodation of freedom. Yet what is the nature of this relationship, and how exactly does freedom determine the shape of liberty? What are the components of the concept of liberty that accommodate the demands of freedom? Does it have to do only with the negative rights, the freedom *from*, or certain positive rights, the freedom *to*, should also be upheld by the society in order to enable meaningful freedom of its members?

In this paper, I will argue that in order to accommodate the demands of freedom, to enable its meaningful expression, political liberty must not only protect the individual from undue limitations but also provide one with the means necessary to carry out free choices. Without such means, freedom will remain merely formal, lacking any essential content. Freedom, seen as the ability to choose between alternatives, can be expressed only in its social circumstances. Therefore, the means the individual has at her disposal to exercise free choice should match the character of the society in which the individual functions, and specifically its level of sophistication.

In order to advance the argument, I will first clarify the concepts of freedom and liberty. This is necessary as the terms are frequently confused, and there is a considerable disagreement regarding the meaning of the conception of freedom. Specifically, I will suggest viewing freedom as the capacity to make choices between alternatives. Then we can examine what features the societal arrangement, i.e., liberty, should possess in order to accommodate freedom, emphasizing the latter's formal and substantial aspects. In doing so, I will analyze the different levels of societal liberty as accommodating freedom along the lines of the Hegelian framework, namely as freedom in itself, for itself, and in- and for itself. I will also look at freedom from the perspective of the information theory. I will argue that we need to recognize the special role of knowledge, and hence the role of liberty in relation to knowledge in order to enable meaningful freedom. Following that, we can explore whether this special role of knowledge might give rise to certain political rights.

2. Clarification of concepts: The capacity to choose vs. its social accommodation

The terms 'freedom' and 'liberty' are frequently used as synonymous, yet a distinction is made whenever needed between freedom as in "free will" and freedom as in "free speech."[1] In the first case, 'freedom' refers to human capacity for self-determination, while in the second — to social arrangement that is related to this capacity. The hitherto most thorough review of the literature on freedom undertaken by Mortimer Adler and his colleagues (Adler 1973) also suggests that two distinct aspects of freedom: freedom as individual's capacity, Adler's *natural* freedom (Ibid. 93, 107, and elsewhere), and freedom as a societal response to this capacity, either in the form of limits imposed on the individual or individual's self-development alongside and against those limits — Adler's *circumstantial* and *acquired* views of freedom (Ibid.). Thus, it seems that the use of distinct terms to denote these two concepts is well warranted. In what follows, I will use 'freedom' to refer to the assumed capacity of the individual, and 'liberty' — to the aspects of societal organization related to this capacity.

2.1 Freedom: Individual's capacity to choose between alternatives[2]

The minimal conceptual essence of freedom would have a number of aspects. First, it must be individual, or accord with the boundary between the self — however we define it — and the outside world. Without that, the notion of liberty would not hold: there is no influence of circumstances that impact individual's freedom without the boundary between the individual and the external world. Second, it should account for the influence of constraints against which the individual will be asserting the degree of freedom attainted. Without such constraints, the notion of acquiring a degree of freedom is meaningless.[3] It is necessary to note that the constraints might vary greatly: these can be solitary confinement and fear of persecution or public mockery, as well as desires of the flesh and temptations of prejudiced thinking. Third, and the most important, freedom must include a sort of causality that is different from the one that is usually described as natural causality, i.e., causality that can be described by observable laws of nature. The essence of freedom is the causality through the self that is not completely determinable by the natural world, usually referred to as self-determination.

In order to meet these criteria we can conceptualize freedom as *ability to choose between alternatives*. This ability can be seen as innate — in fact, it must

1 See, for example, Mill's *On Liberty* (Mill 1859/2010, Chapter I, para. 12-13).

2 J. Melvin Woody in *Freedom*'s *Embrace* arrives at a characterization of freedom similar to the one given in this section through analyzing what the hypothesis of freedom would entail so it can stand the test of human experience (Woody 1998; see 19-20 for the brief outline of the approach and Parts I, pp21-64 and III, 129-228 for the development of the argument).

3 An argument for the impossibility of absolute freedom can be found in Woody 1998, 85-112. For our purposes, it would suffice to say that there would be no need for a conception of freedom if there were no constraints upon it: the Emerald City needs no color word for green.

be seem as innate, as without it the discussion of freedom would be meaningless. Yet it does not have to have any specific content, as such content would be, at least potentially, determinable by a variety of factors.

An important aspect of the analysis of freedom as the capacity to choose, thoroughly addressed by Locke, is identifying the factors that constrain available choices, and specifically the role of understanding (perception/thinking) in our judgments. Not only physical constraints are at play here, our thoughts also influence our choices. If one has little idea about traffic signs, his choice might well be influenced — at times, at a considerable costs and inconvenience — by this lack of knowledge. A person who is convinced that "theory" signifies something vague and uncertain is likely to choose a different course of action when called to decide on the matters of nature than a person who is familiar with how science functions. As Locke puts it, "without understanding, liberty (if it could be) would signify nothing [...] he that is at liberty to ramble in perfect darkness, what is his liberty better than if he were driven up and down as a bubble by the force of the wind?" (Locke 1691/1959, §69, 361). I will try to show that wandering in the twilight is also quite problematic.

2.2 Liberty: Societal accommodation of freedom

If freedom is a capacity natural to human beings, society ought to address it: since society is comprised of individuals, their constitution, both mental and physical, is of an essence for social arrangements.[4] There will be, thus, a societal accommodation of freedom, or liberty.

Liberty would be established to allow for the optimal expression of freedom to the extent possible within the constraints imposed by the necessities of living in a society. Hence, a discussion on liberty can proceed in two planes: the constraints on individual freedom and the enablement of its development. The aspect of constraints is reflected in the circumstantial view of freedom (Adler 1973, 93, 107, and elsewhere), as well as in the concept of negative liberty explicated by Berlin in his *Two Concepts of Liberty* (Berlin 1969/2002b). It is also clear that the notion of negative liberty accommodates the conception of freedom as choice that, following Woody, has been proposed in this paper. If the ability to choose between alternatives is natural to individuals, protecting this choice seems to be natural to societies. Yet liberty that focuses on the enablement of freedom, liberty that is equally necessary for accommodating freedom, does not seem to be adequately addressed in the literature.

Berlin, most certainly informed by the social upheavals to which he had been a witness since he was seven years old, opposes to negative freedom the notion of *positive* freedom (Berlin 1969/2002b, 178), or freedom *to* (Berlin 1998/2002, 326).

4 This seems to be the assumption behind political philosophy since the times of Plato's *Republic.* Plato's view of people as impressible by stories led him to impose censorship in his ideal city (Plato, 386[a]-389[a] in Plato 1997, 1022-1026). Similarly, Locke in *Two Treatises of Government* argues against the views of his opponent Filmer regarding whether men are born free or not; both see in it the basis for how the government is to be organized (see, for example, Locke 1689/1988, Book I, §2, 142; Book II, §4, 269; and elsewhere).

Positive freedom starts with the question "Who governs me?" as opposed to "What are the limits of the control others can exercise over my choices?" that is central to negative freedom. The development of this view, per Berlin's observations, inevitably leads us to base our judgments on the conception of the self. Bhagavad Gita's detachment, Aristotle's virtues and Stoics' self-discipline (see review in Adler 1973), and Kant's discussion on autonomy of the will as opposed to the heteronomy of desires (Kant 1785/1998 and Kant 1797/1996) can serve as examples of this approach. These sources seem to stress individual development rather than social arrangements; even educational treatises written within this tradition do not rise to the level of society at large.[5] And yet many modern political movements act according to the notion of positive liberty: furthering the "real" freedom of the "real" self, be it the rational nature, nation, soul, or another similar concept that is claiming to be truer and higher than the perishable flesh and ignorant calculations of the individual. The examples Berlin repeatedly brings are those of inquisition, communism, fascism, and nationalism of different sorts; in our times we can add to the list religious fanaticism not organized hierarchically and, peculiarly enough, the almost-religious belief in the highest truth of the mysteriously invisible hand of the free market, that among its ardent proponents seem to evolve from a metaphor into a value in itself. It is clear that this notion of positive freedom can be used to limit and crash the mere notion of freedom as the realization of the ability to choose among alternatives.[6]

Is the exclusion of positive freedom from the domain of liberty warranted? Berlin has been criticized extensively for this move. The most interesting criticism, in the context of this paper, comes from MacCallum (MacCallum 1967). He claims that freedom involves a triadic relation: it is always *of* somebody, *from* something, and *to* do something (Ibid. 314). In other words, freedom always has an agent, its subject; a limit, at least a potential one; and an object, its goal. What Berlin refers to as a negative freedom, claims MacCallum, is the freedom *from*, yet it is meaningless without the freedom *to*, Berlin's positive freedom: what does freedom from censorship mean to an agent who is not about to read or write any books? The recent communitarian treatment of positive freedom adds more substance to this claim. Without the rich context of culture, society, and history, with only the most basic and simple forms of decision making at her disposal, the individual isn't free but is a rather shallow and narrow atomistic entity that cannot meaningfully carry out the choices.[7]

Berlin's response seems to be quite convincing. There is a sense in which breaking from the chains of oppression has meaning without any particular course of action planned — the sense of being able to choose without repercussions, i.e., freely (Berlin 1969/2002a, 36n; Berlin 1998/2002, 326). The desire of a person to be free

5 See, for example, Kant's *Education* (Kant 1803/1964), the primary focus of which is the individual.

6 Berlin most famously notes it; his selection of examples adds historical validity to this view. See, for example, Berlin 1998/2002, 328.

7 See, for example, Taylor 1992, 40-41 and throughout Chapter I. Other prominent communitarians are surveyed in Etzioni 1998.

in the sense of being able to carry out his choices can be compared to the desire of a deaf person to regain the ability to hear: the question of *to*, i.e., what specifically he wishes to hear as a reason to regain hearing, would be immaterial.

And yet MacCallum's critique and the communitarian concerns, as many other responses to Berlin's *Two Concepts of Liberty*,[8] point out an important corollary of having negative liberty as the sole meaning of liberty, important specifically in the context of seeing liberty as freedom's social accommodation. If society's role is only to make sure that freedom as choice is not limited beyond what is necessary for its own maintenance, where would the contents for this capacity, the contents necessary to distinguish the alternatives of choice and make informed choices, come from? Freedom of the press seem somewhat problematic, to say the least, in a society where the vast majority is illiterate and Internet is inaccessible. Mere literacy would not be enough either if the choices to be made require understanding of advanced concepts and the ability to analyze complex data: consider the decision for or against coal-powered plants, when done by people whose knowledge of natural sciences is vague. Moreover, noting that by making choices freedom can shape itself through setting the circumstances for its own future application, we will arrive at the understanding that specific contents can lead it to limit or deny itself. Arguably, naïve fellows taught to respect authority, exposed to nicely packaged ideas of absolutism and denied access to alternative concepts, whether by censorship or by the lack of acquired ability to follow sophisticated argument, might well deny their natural freedom. Mere negative liberty, which, in the context of our distinction between freedom and liberty, can be better deemed *protective liberty*, would not be enough to accommodate the inherent human capacity for choice at the societal level — it will leave it empty of adequate content.[9]

From here, society's role in accommodating freedom cannot be confined to making sure that no unnecessary limits are imposed on individuals' capacity to choose their course of action. Nor would it have much to do with Berlin's idea of positive freedom. In order to provide an adequate response to this basic element of human nature, society has to ensure that the alternatives of choice are present and accessible, and that the individual is equipped with what is needed to make rational choices — that the form of freedom receives content over which it can be exercised. This can be deemed *enabling* liberty, and as such it complements the *protective* (negative) liberty.

3. Freedom and liberty: The relationship of enablement

The concept of freedom as a capacity to choose between alternatives has a number of consequences. First, it requires constraints. A choice can be made only when we have a number of specific alternative courses of action; the fact that these are specific alternatives and that there is a finite number of them both enables and limits

8 See survey in Berlin 1969/2002a.

9 Putterman also argues for taking into consideration the content of freedom while analyzing Berlin's views (Putterman 2006, 421, 425, 438).

our choice. From here, the number and the quality of choices would be in a positive correlation with the degree of freedom, yet this degree will never be absolute. Second, to realize itself, freedom needs access to alternatives and should be capable of making choices. If no alternatives are available, choice is impossible. If the agent is incapable of making the choice, it is equally impossible. Yet modeling freedom after a subject in a psychological experiment who is requested to choose between three alternatives regarding which she has all the relevant information would be highly misleading; it is what Taleb deemed *ludic fallacy*, seeing human interaction with all its complexities as a sort of simple game with well-defined rules (Taleb 2007). Having access to information about alternatives, as well as the ability to process such information and understand the consequences of choice are necessary if we are to talk about real choices made in the complex world of any human society, ancient Greece as modern Denmark. And this is where enabling liberty becomes relevant.

In order to clarify what enabling liberty should consist of, it would be beneficial first to briefly address its boundaries. Enabling liberty cannot provide goals for choice, neither can it guide toward preferring one choice or group of choices over another. Equality, morality, and other values can do just that. However, none of these accommodates choice *qua* choice. A chess example would help here, as this game seems complex enough to exemplify issues from the world of human interaction. Teaching somebody to play chess would entail familiarizing her with the rules of the game, i.e., the moves pieces are allowed to make — the alternatives of choice. Theoretically, this is enough, as everything else can be derived analytically from the rules. However, if our neophyte is to confront an opponent within a week, mere communication of rules is far from being enough. Forks, pins, defense and attack strategies, etc. would be of real value and will certainly enhance one's ability to play a meaningful game of chess. However, this would not be enough either, as alone it will not help the player to evaluate the options and choose the best one. Criteria for appraising alternative moves and selecting the best one, as well as guidance for applying these criteria and formulating new ones, would be of high value. All this together will make a good chess player without pushing her in one specific direction — enabling rather than directing. While real life-choices are much more complex and consequential than chess moves, the example does demonstrate the three main elements necessary for enabling liberty: access to alternatives, tools for the analysis of alternatives, and methods for developing criteria for evaluating the strategies of choice and forming new ones.

These three elements of enabling liberty can be seen as related to *knowledge*: knowing *what* the alternatives are, knowing what they *mean*, and knowing how to *evaluate* them. Yet before these elements are analyzed as progressing levels of freedom, it is necessary to recognize knowledge's unique role in its enablement. While mentioned by Berlin, Adler, Woody, Sen (Sen 1999), and others, knowledge has never been assigned a unique place in the freedom discourse. However, it seems that its role is distinct from all other enabling factors. These factors, e.g., physical conditions, can be hardly overestimated in their importance, but none of them seems to be *necessary* for carrying out free choices. It can be argued, as Berlin does in

Two Concepts of Liberty and elsewhere, that the freedom of the Stoic is limited—but it still can be validly called freedom. The Stoic deliberately limits his choices to avoid constraints, thus proving that many important factors, including, most notably, physical and legal conditions, can be discarded when one restricts his realm of choice. Yet without the knowledge of the alternatives the choice is not only difficult—it is impossible. Hence, knowledge constitutes a *necessary* pre-condition for the realization of freedom: knowledge provides it with contents, without which freedom cannot be carried out in the world, as minimal as it might be. Not knowing what the options are equals to not being able to choose.

4. Three levels of liberty

The three levels of knowledge mentioned above can be looked upon as the development, or unfolding, of freedom in the Hegelian sense (Hegel 1807/1977), as well as in terms of the information theory. Analyzing the development of freedom along Hegelian lines enables tracing the essential connection between social liberty and individual freedom. Looking at the levels of knowledge through the lenses of the information theory allows better understanding of the possible ways of accommodating freedom in social practice.

4.1 Access to alternatives: freedom *in itself* as data

The most basic level of enabled freedom is access to alternatives. Without accessible alternatives of choice, freedom forever remains merely formal, unrealizable capacity. Some alternatives are accessible to us by virtue of our human nature—the classical example that seems to occur to any philosopher discussing the subject is raising a hand or not doing so. This, however, means little for the purposes of human freedom, as similar alternatives are accessible to any mammal. To understand this level of freedom better, we can see it addressing freedom *as potential*: access to alternatives provides options that are necessary for implementing choice, yet no more than that.

Having alternatives accessible is a necessary condition for carrying out free action—yet by no means sufficient. It is also necessary that the agent *understands* the alternatives of choice *as* alternatives of choice. Without this understanding, possible courses of action remain alternatives only *in themselves (an sich)*, much as an embryo that is "*in itself* a human being, [but] it is not so *for itself*" (Hegel 1807/1977, §21, 12)—and hence the freedom of choice remains unrealized. Following Hegel, we can think of a slave who has all the necessary means for the insurrection accessible, yet does not perceive these *as* means for the insurrection since he does not see himself as free to revolt—and hence, in our terminology, has mere access to alternatives for choice yet does not see them as such.[10]

In terms of the societal accommodation of freedom, enabling liberty here ought to make sure that the alternatives are present and accessible. This is akin

10 See Hegel's discussion on freedom in Oriental society (as he understood it, of course) in Hegel 1837/1953, Ch. 1. The Idea of Freedom, 23-24.

to providing data, where data is understood in terms of the information theory as entities that can potentially be interpreted (Floridi 2010, 23-24). Such role can translate into a wide spectrum of social action. On the one end stand voting rights and other ways to express political will, infrastructure for the freedom of movement, etc. On the other — establishing and maintaining public libraries, ensuring that media is not monopolized — or, better, stays non-monopolizable,[11] and providing Internet access for all. All these offer avenues for implementing choices: the first step in enabling freedom. However, data is meaningless when the means of its interpretation are lacking — this is easy to see if we think of a text in a language unknown to us. Similarly, libraries are useless to the illiterate, and the web — to those who have neither means nor the skills necessary to make meaningful use of it.

4.2 Understanding alternatives: freedom *for us* as information

Information stands for well-formed meaningful data (Floridi 2010, 2). As such, it is qualitatively different from data which by itself is neither well-formed not meaningful for its users. Similarly, at the second level of freedom's enablement the alternative courses of action are not merely accessible but also understood by the agents as possible ways to act. Here the alternatives are *for me (für mich)*, I can understand them as something that can be pursued. This is the second necessary step toward freedom's realization, as it is impossible to choose something not seen as a possible alternative.

Understanding alternatives as such constitutes significant progress when compared to the mere access to alternatives. There is a qualitative difference between having a legal ability to vote and knowing that you can vote: when, where, and how; having a library in town — and knowing that it is available for you; having access to the Internet — and using it. At the first glance, it might seem that having access to the alternatives of choice and understanding them as possible courses of action is enough to realize freedom. Yet it is not.

Complex alternatives require more than mere encounter to understand what their value is, they necessitate more than just having information to make a meaningful choice. Without means to analyze the information, the "web of mutual relations," information has little meaning; without such relations "you are left with a pile of truths or a random list of bits of information that cannot help to make sense of the reality they seek to address" (Ibd., 51). One might know that Nietzsche wrote a poem named *Vereinsamt*, and even be able to get its full text on the computer screen in seconds and read it; and yet not being accustomed to reading poetry will make the prospect of enjoying it impossible, and the choice –meaningless. One might know how to vote in general elections, have a full right to do so, and not fear any repercussions; but if she has no ways of understanding — not merely reading, but understanding — the programs of the candidates and the possible consequences of these programs, voting loses its meaning as a choice made between alternatives and becomes an exercise in a skewed game of chance. A person can have full

11 Its seems like we are witnessing the creation of non-monopolizable media through the combination of Internet technology and its skillful use by millions of people around the world.

access to all publicly available information on global warming, but if his chemistry education was limited to one semester of re-hashing definitions from the textbook, as a result of which he perceives natural sciences to be a sort of opinionating regarding things that has little to do with the real world, he will not be able to appreciate the information and make meaningful choices in regards to it. One might be aware of two possible choices, yet not even fathom that the whole situation can be re-conceptualized by applying a new paradigm to it, thus increasing the number of alternative actions.

Modern liberal democracies have political freedoms enshrined as laws of the land and thus protective liberty in place. They also usually succeed in providing data to their citizens and equipping them with the way to turn it into information, for example, through public libraries and literacy. Yet, as shown above, this is not enough for making meaningful choices on complex matters and hence it does not properly accommodate freedom. The situation is akin to Hegel's Greeks who were conscious of freedom, yet did not see "man as such" as free (Hegel 1830/1971, §482, 239; Hegel 1837/1953, 23). Their freedom, consequently, was partial and accidental, where one is seen free *thanks to* something external, e.g., place of birth, rather than her own human nature. Freedom here is not the essence of life but rather one bit of reality among many—just like alternatives of choice are "bits and pieces" of information, separate from each other and not integral in their role *as* alternatives of choice to the rest of the fabric of life.

Charles Taylor analyzes this situation as resulting from the dismantling of the traditional society, where every person was placed in a specific station in life, with its roles and responsibilities, accompanied by a full repertoire of knowledge necessary to living his life. These certainly were restrictive, and yet "at the same time as they restricted us, these orders gave meaning to the world and to the activities of social life" (Taylor 1992, 3). When the traditional society is replaced by the mere freedom *for us*, the place of meaning remains void, our reasoning cannot be but merely instrumental (Ibid., 8-9; see also Lyotard 1984), and our freedom—only partial and not fully human.

4.3 Evaluating alternatives: freedom *in and for itself* as knowledge

According to Hegel, the ultimate realization of Spirit (*Geist*) is in its being *in-and-for-itself (An- und Fürsichseiende)*, where it realizes that *in-itself* and *for-itself* are two moments of its existence (Hegel 1807/1977, §804, 490). Applied to freedom, this is reflected in a human being who realizes herself as free *qua* human being, realizes freedom as the nature of humanity (Ibid., §482, 240). It entails a vantage point from which the ability to choose and the alternatives of choice are seen as part of one realm—the realm of freedom.

This leads us to knowledge. In terms of the information theory, knowledge would refer to a "web of mutual relations that allow one part of it to account for another [... in which] information starts providing that overall view of the world which we associate with the best of our epistemic efforts" (Floridi 2010, 51). Applied to freedom, knowledge would mean having a rich context against which the alternatives can be evaluated. The context here includes methods for evaluating

the alternatives as well as means to formulate these methods; methods for arranging information in such a way that it would make proper historical or otherwise factual background for the choices under consideration; being able to ask a number of meta-informational questions — questions about the relevance, usefulness, reliability, possible interpretations, level of details, and veracity of information (cf. Floridi 2010, 48, 52). Knowledge here is not merely understanding something but comprehending its meaning.

Knowledge goes beyond the ability to manipulate information, beyond comparing *n* alternatives over scale *s*. The full meaning of enabling freedom is in providing the resources that enable and empower human beings to *think*: to raise questions of value and meaning and have the resources needed to answer them as such. Freedom empowered by knowledge is the freedom to read a poem while being able to put it in the context that allows the reader to appreciate it. It means understanding that evolution is not a subject of belief but, as scientific theory, of support or falsification by empirical observations. It entails not only knowing that different parties are soliciting votes and that each one of them has certain ideas and agenda, but being able to understand these agendas in their historical context, see them as elements of a particular political system that accords specific privileges to elected officials, and understanding what is the depth of their impact on the course the country is about to take after the elections. Moreover, the rich context here increases the freedom not only by giving it meaning but also by enabling and empowering the agent to look not merely at the available alternatives and evaluate them but also to evaluate the paradigm with which she construes the situation of choice and, if desired, come up with the new one. The context here enables a meta-choice, a choice of the strategy of choice, as opposed to acting within the model given by habit or tradition. Here lies the principal difference between the suggested view of liberty as freedom's enabler and the solutions usually proposed by the communitarian thinking.

The argument proposed above leads to the considerations that seem to be behind the ancient Greek ideal of a well-rounded person and the Confucian *chün-tzu* (Confucius & Waley 1989, specifically Book II). However, this level of freedom's unfolding in the social realm is yet to be attained by liberal democracies. Moreover — it seems that the development of education and the public discourse after World War II have been moving in the opposite direction (Lyotard 1984). Yann Martel reflected on similar issues during a session of Canadian parliament to which he was invited:

> [...] to think of the arts as mere entertainment to be indulged in after the serious business of life, that — in conjunction with retooling education so that it centres on the *teaching of employable skills rather than the creating of thinking citizens* [italics mine – MY] — is to engineer souls that are post-historical, post-literate and pre-robotic; that is, blank souls wired to be unfulfilled and susceptible to conformism at its worst — intolerance and totalitarianism — because incapable of thinking for themselves and vowed

to a life of frustrated serfdom at the service of the feudal lords of profit.[12]

In order to accommodate freedom, liberty is to make sure that knowledge is fostered and developed by the educational system. This, not in order to answer the call of the "real self," as with Berlin's positive freedom, but to allow for the real, meaningful, human choice.

5. Conclusion: From meaningful freedom to political rights

The analysis of freedom attempted above leads to the conclusion that there are two aspects of the societal accommodation of freedom. The first focuses on making sure that the interference with the choice made by individuals is minimal; rather than the traditional name of negative liberty, it can be better called *protective* liberty to reflect its meaning. The second aspect is as necessary to make freedom shine as the first one. Since the capacity to choose is formal, it needs contents to be realized, contents about which the decisions are to be made. *Enabling* liberty comes to make sure that the individual can make meaningful choices. As such, it needs to address three levels of freedom as choice making — and these levels seem to be translatable into specific political rights.

First, in order to enable freedom, access to the alternatives among which the choice will be made should be provided. In our society it can be translated into the protection of access to information and provision of such access — and to the respective political right to information. This right, though, cannot remain a mere abstraction: just like the right to travel translates into a transportation infrastructure and legal arrangements, the right to access information ought to be expressed in accessible and adequate informational infrastructure. Moreover, it does not merely mean access to data banks but also access to processed and organized sources that will help in processing the information. This idea is not as new as it might sound, and it might be traced to Mill's view of society's role as the repository of information related to social experiments (Mill 1859/2010, Chapter V, para. 18). Public libraries were the beginning of this process, which today seems to require unimpeded web access, equality of Internet content protected by law, and protecting sources like Wikipedia from the encroachment of interest groups and political players. Moreover, it would seem imperative for societies committed to the ideals of freedom to disseminate knowledge and information that will help people in countries where access to information is restricted in making informed and knowledgeable choices; this is principally different from governmental propaganda, as the dissemination of information meant here is not committed to this or that particular position.

The second political right that comes to mind is the right to education that provides knowledge, as opposed to one that merely equips students with skills. As

12 The quotation is taken from the website where Yann Martel is tracing his project of sending a different book to the Prime Minister of Canada every two weeks. As of time of writing this paper, he has mailed 101 books yet received no significant feedback.

primary means of enabling the members of society to construct alternative courses of action and evaluate them, education, in order to answer the call of freedom, should match the complexity of choices the individuals are expected to face. The more complex society is, the higher should be the quality of education. The alternatives individuals confront in modern society develop and change rapidly. Therefore, mere supply of facts and specific criteria for evaluating alternatives would not be adequate for today's world—hence the need to teach *how to learn*, or to provide the individual with the means to acquire, produce, and disseminate new knowledge; and to be aware of herself as being able and needing to do that.

To summarize, if we are to accommodate natural freedom at the societal level, it is not enough to protect it from interference. While freedom without protection is incapable, without enablement it lacks content; without protective liberty freedom can disappear, yet without enabling liberty it might well lose any meaning.

References

Adler, M. J. (1973). *The idea of freedom*. Westport, Conn.,: Greenwood Press.

Berlin, I. (1969/2002a). Introduction. In H. Hardy (Ed.), *Liberty: incorporating four essays on liberty* (pp. 323-328). Oxford: Oxford University Press.

Berlin, I. (1969/2002b). Two Concepts of Liberty. In H. Hardy (Ed.), *Liberty: incorporating four essays on liberty* (pp. 166-217). Oxford: Oxford University Press.

Berlin, I. (1998/2002). Final Retrospect. In H. Hardy (Ed.), *Liberty: incorporating four essays on liberty* (pp. 323-328). Oxford: Oxford University Press.

Confucius, & Waley, A. (1989). *The Analects of Confucius* (1st Vintage Books ed.). New York: Vintage Books.

Etzioni, A. (1998). *The essential communitarian reader*. Lanham, Md.: Rowman & Littlefield.

Floridi, L. (2010). *Information: a very short introduction*. New York, NY: Oxford University Press.

Hegel, G. W. F. (1807/1977). The Phenomenology of Spirit.

Hegel, G. W. F. (1830/1971). *Hegel*'s *Philosophy of mind: being part three of the 'Encyclopaedia of the philosophical sciences' (1830)*. Oxford,: Clarendon Press.

Hegel, G. W. F. (1837/1953). *Reason in history, a general introduction to the philosophy of history*. Indianapolis,: Bobbs-Merrill.

Kant, I. (1785/1998). Groundwork of the metaphysics of morals [Electronic reproduction]. Cambridge, U.K. ; New York: Cambridge University Press,.

Kant, I. (1797/1996). *The metaphysics of morals*. New York: Cambridge University Press.

Kant, I. (1803/1964). *Education*. Ann Arbor]: University of Michigan Press.

Locke, J. (1689/1988). *Two treatises of government* (Student ed.). Cambridge Cambridgeshire ; New York: Cambridge University Press.

Locke, J. (1691/1959). *An essay concerning human understanding*. New York,: Dover Publications.

Lyotard, J.-F. (1984). *The postmodern condition: a report on knowledge*. Minneapolis:

University of Minnesota Press.
MacCallum, G. C., Jr. (1967). Negative and Positive Freedom. *The Philosophical Review, 76*(3), 312-334.
Mill, J. S. (1859/2010). On LibertyOn Liberty and Other Essays (Amazon Kindle ed.): digireads.com.
Plato. (1997). *Complete works* (J. M. Cooper & D. S. Hutchinson, Trans.). Indianapolis, Ind.: Hackett Pub.
Putterman, T. L. (2006). Berlin's Two Concepts of Liberty A Reassessment and Revision. *Polity, 38*(3), 416-446.
Sen, A. (1999). *Development as freedom* (1st. ed.). New York: Knopf.
Taleb, N. (2007). *The black swan: the impact of the highly improbable* (1st ed.). New York: Random House.
Taylor, C. (1992). *The ethics of authenticity*. Cambridge, MA: Harvard University Press.
Woody, J. M. (1998). *Freedom*'s *embrace*. University Park, Pa.: Pennsylvania State University Press.

The Professional Morality of the Documentary Filmmaker

Wu-Tso LIN

1. Why need professional morality of documentary?

What is documentary? Can documentary film tell truth? How it tell us the truth? These questions related to the working mindset of documentary filmmakers. While looking for an appropriate documentary topic, the first consideration came to the director's mind and inspire his passion as well as to search out a topic other people care about. But what do people care about?

For example, to begin with humanitarian issues are at the forefront of the global consciousness, such as human being worrying about the future and maintaining a sustainable way of life, world peace and human rights (such as Palestine), social experiments, gender equality, racial equality, the balanced distribution of wealth and resources, moving personal experiences and interviews with specialists, historical background. Secondly political issues are hotly debated in most corners of the world, and the director often to film the sensitive situation between different cultural believes, political inclinations and seek to find possible solutions to political tension. For example, the issue of Taiwan's 228 incident, the bloody crackdown against the local Taiwanese people that occurred in 1949 is a little known event in Taiwanese history that would be an ideal topic for the documentary. Another prime example would be the political suppression of aboriginal culture all throughout the American continent and land grabbing by government and big business causing the rapid depletion of the rainforest. Thirdly, environmental issues in this day and age are pressing concerns, such as greenhouse gases causing rising temperatures, depleting our nonrenewable resources, the use of biological and chemical weapons, the dangers surrounding nuclear power facilities, continue finding sustainable and reusable energy, current environmental pollution events such as the island of plastic garbage in the Pacific or nuclear accident of Fukushima Daiichi nuclear power station in Japan. Finally, such as financial issues, any evidence to support your case, conspiracy theory, status quo, hidden camera(privacy rights), above of all are the topics which the documentary filmmakers love to process.

We have to care about the moral problem before our documentary filming. For any kind of professions has internally guidelines of practice that members of the profession must follow, to prevent exploitation of the customer and preserve the integrity of the profession. This is not only for the benefit of the customer but also the benefit of those belonging to the profession. This is so called morality that allow the profession to define a standard of conduct and ensure that individual practitioners meet this standard, by disciplining them from the professional body if they do not practice accordingly. This allows those professionals who act with conscience to practice in the knowledge that they will not be undermined commercially by those

who have fewer ethical qualms. It also maintains the public's trust in the profession, encouraging the public to continue seeking their services.

Here we will be examining the professional morality of documentary film. The issues in Western film studies in recent years have gradually become a school of learning. Currently, there are a few books written by the two philosophers: Emmanuel Levinas and Jacques Derrida as well as Lacan's psychoanalytic science, feminist thought, postcolonial studies and queer theory explore the ethical issues of films (including documentaries). In addition, regarding the professional morality (subject moral rights) of photography, film, television, and digital media, the West has also published research results through some of the works that carry the tripod (such as Gross, Katz & Ruby, eds. Image Ethics). In researching the ethnography and sociology in the field of film and photography, the International Visual Sociology Association (IVSA) in 2009 set out their own standards for governing and research ethics. The American Anthropological Association (AAA), the British Social Anthropology Society (BSAS) and Australian Anthropological Society (AAS) have also established and published their own versions of the code of ethics. There are other books in the field of anthropology, archaeology and sociology, together with many papers and articles that have published their finding in this field. Recently, the Ethnographic Images Society with the National Science Council in Taiwan have promoted the HRPP research ethical norms implementation plan, and developed the "Taiwan Institute of ethnographic video research ethics" (draft) for practitioners of the Association, as well as the domestic audience.

As professional documentary filmmakers we are accountable, and we have to make time for moral reflection. Sometimes, it is difficult to find time to consider a moral position during filming. However, as a documentary filmmaker, we appreciate how tough it is when considering these issues. Under these conditions, we need more time and a special applied philosophy to face the situations and overcome the difficulties. This essay begins with the legal limits imposed on the filmmakers and explores some methods and principles in dealing with moral considerations and ethical dilemmas at every stage of the documentary filmmaking process.

2. Documentary work and the law

Every profession has their special form of ethics, named professional morality. For example, when a doctor treats his patients, he should treat them as own family. More than that, a doctor has to inform patients of all possible treatments and healing methods that he could use in a specific situation.

We can, therefore, understand that professional morality exists within every kind of profession. We need to follow our inherent moral standards in order to fulfill our responsibilities to the people we encounter in the workplace so as to safeguard their rights.

The categories of professional morality belong to a person's inner consciousness and autonomy. Without this sense of professional morality consequential harm may be inflicted. An end result of this may even include legal

action in serious cases.

Although principles on professional morality within documentary film making can often unclear, there are times when their effects are clearly felt. For example, the documentary filmmakers must, on a daily basis, consider the legal implications of their film productions. There are limits that protect the privacy of individual under the law. Due to the interference of filmmakers in the past, laws have been enacted to prevent any intrusion into the private lives of the public.

For example, in 1890, two Boston lawyers, Samuel Warren and Louis Brandeis published an article in the Harvard Law Review which has come to be regarded as the outstanding example of the influence of legal periodicals upon the American law (Prosser, 1960:383). The article was called "The Right to Privacy," and it was motivated by the prying of the press into the social affairs of Mrs. Warren and her friends. Warren and Brandeis argued for the existence of specific right to privacy which was implicit in other recognized rights, "and they contended that the growing abuses of the press made a remedy upon such a distinct ground essential to the protection of mental distress" (Prosser, 1960:384). In the years since the article was published, the right to privacy has been invoked by plaintiffs, contested by defendants, and often accepted by judges and juries, in a bewildering array of cases (Larry Gross, John Katz, Jay Ruby eds., 1998).

The most influential analysis of privacy implication is that provided by William Prosser who declared that there were four categories of privacy invasion. These four categories are comprised of intrusion, embarrassment, false light and appropriation. All these categories that involve failures of professional morality should be avoided by the documentary photographer or filmmaker.

3. Documentary work and professional morality

Laws that protect an individual's privacy are useful tools and serve as concrete guidelines for the filmmakers to follow. We can, at the same time, provide ourselves with proactive and positive working ethics. We can hope to achieve a high level of professional morality when faced with moral and ethical situations. They can help as when tackling some of the fundamental questions of our profession: How should we work? What should we do in a particular moral dilemma? With understood and accepted moral principles the effectiveness of our decisions making process with regards to professional morality can be greatly improved.

We can say that morality represents a higher level of doctrine in profession in laws. This is because morality constitutes the inner universal law inherent in human beings. Morality can use our rational intuition to be applied within ones' workplace. The practical ethics that the Greek philosopher Aristotle espoused states that ethical knowledge is not only a theoretical knowledge, but also that a person must have "experience of the actions in life" and have been "brought up in fine habits" to become good (NE 1095a3 and b5). For a person to become virtuous, he cannot simply study what virtue is, but must actually do virtuous things. To sum up, ethics is a virtue, a kind of morality which encourages a person to reach his highest goal

in work and in his thought. Another philosopher, Britain G.E. Moore and ancient Chinese philosopher Mencius both believed that morality could not be taught and analyzed as it is a natural intuitive and instinctive response. The discussions above all also have established a kind of philosophy of film (James Donald and Michael Renov eds., 2008).

We can apply these theories on morality discussed above and apply them to the work of the artistic creators within documentary filmmaking.

As we have previously examined the abstract notions of this moral philosophy, we therefore need to deal with the more concrete and practical ways with which to apply these principles to the workplace. In doing so, we must avoid perceiving the result of the outcome of our work as simply "good" and "bad". It is also important to carefully analyze the whole process of filmmaking. We will now construct a framework with which the filmmakers can follow so as to create a common understanding shared by those working within this field. In wanting to layout applicable principles of professional morality, here are seven practical criterions that can be adopted.

3.1 Never deceive or cheat the participant

In the creative process, the filmmaker should seek to be honest with regards to the aims of the documentary.

An example of this can be found in the filmmaker use of hidden cameras during filming. This kind of behavior demonstrates a serious fault in the film making process. It is not only unethical, but also violates the law. We should therefore avoid to use of hidden cameras or other hidden equipments when recording.

Furthermore, we must ensure that the participant's understanding of their function in the film is clear. Just as we refuse to mislead the audiences, so we must not cheat or deceive the participant or interviewee. Documentary film, commonly referred to as "documentary" or "documentation" differs considerably from narrative or fictional films, even those with stories that are based on real life events. The primary difference is that documentary films are obliged to adhere to the truth without alteration or elaboration. In a word, the biggest difference between documentary and fiction (drama) is the difference between "truth" and "imagination" (Michael, 2008). Fictional film may include the presence of actors, script and high quality filming techniques. Its appearance therefore is artificial. In contract, the essence of the documentary style necessarily requires truth. This style is also referred to as direct cinema or cinéma vérité. The whole process of filmmaking is reliant on the recreation of real event. To summarize this point we can simply suggest that honesty is the only policy in the creation of documentary film.

3.2 Never Change or Modify the Scene of Filming

The documentary filmmaker should never seek to alter the scene of filming to achieve a more convenient or aesthetically pleasing result. If it happens, the audience will be misled. If, for example, you organize or clean the scene it will mean that the relationship between the protagonist and the environment they are in will be changed. The documentary filmmaker should only be a spectator, and observer, and

so the scene of filming should never be changed or modified.

In cases where there is no supporting evidence that provides information of the original scenes these entail special considerations. In this circumstance the documentary filmmaker is oblige to restructure the scenes to create and honest interpretation within the overall context.

3.3 Never film by force and never film with the purpose of earning personal fame and money

A professional documentary filmmaker will not begin the filmmaking process or indeed end it due to reasons of money, personal status, the threat of violence, coercion etc. If the documentary filmmaker wants to retain objectivity and independence, he should position himself away from any temptation or benefit personally from the filmmaking process. Benefits arising from government or large business can be especially corrosive to any objective stand point. Furthermore, the documentary filmmaker must be on guard as to prevent himself becoming a tool of the same groups.

The youth of today may want to earn money or obtain personal fame by making documentary films. The motivation to produce film can be negatively influenced by a desire for winning prizes or the promise of a commanding reputation. When the goal of producing honest and sincere work is tainted by the temptation of personal gain the moral high ground is inevitably lost.

3.4 Never fake the recording, twist the facts or modify the truth. never fragment the scene, change the chronology, alter the conversation or misinterpret an interviewee's intention or expression in the editing process

When editing the documentary in the post-production process, a filmmaker should adhere to the original context and chronology of the scene that is represented in filming. In addition, as the camera is the observer we should refrain from adding anything more to lens.

The exaggeration of meaning or exploitation of the story or characters through any overuse of editing should be really avoided as this too would compromise the integrity of the documentary. An example in this case would be disingenuous TV news stories, that distort the facts reported through the addition of computer generated images or sound. This method is contrary to the nature of the professional news and documentary film production.

3.5 The filmmaker can have a point of view but must avoid using the media to express any radical views or positions

Although documentary films vary significantly in style and point of view, the filmmaker must remain somewhat free of outside influence. A perspective of the facts and information has to be rational. Any radical or emotional views expressed can be dangerous with regards to the films objectivity (Erik, 1974).

Looking at the full spectrum of films within the documentary film genre, we can find more conventional documentaries, such as the film "*In Search of Mozart*",

which uses moving and still images to introduce you to their subject. In this case, it is the life and music of young Wolfgang Amadeus Mozart that is examined. At the other end of the spectrum, we can find the works of experimental documentary filmmakers such as Abigail Child, whose film "*Mayhem*" is a highly personal expressive piece that uses unexpected juxtapositions of images.

Some documentary films may feature their own filmmakers, such as Michael Moore in his film "*Sicko*". Other filmmakers may use reconstructions to depict a scene when actual footage isn't available. Michael Winterbottom's "*Road to Guantanamo*" is an example of that technique.

3.6 The issue of privacy

A documentary film can only be completed when both the filmmaker and participant agree on its publishing. If the participant requests that their identity be hidden this must be respected. For legal and ethical reasons the protection of privacy for participants is essential.

The future of documentary films relies on the trust of participants and their willingness to come forward. As filmmakers we must honour this trust so as to encourage participation in future projects. In cases of the use of multiple participants a broadcast agreement can be made and signed to ensure that trust and confidentiality is respected by other participants involved in the film.

3.7 Regarding the participants as partners or co-authors.

The production of documentary film stands in stark contrast to arts including painting, sculpture, pottery, photography, dance, literature, or music. The essence of documentary filmmaking relies on the participation of both interviewee and interviewer. It is for this reason that we must recognize both parties as being active partners or co-authors.

Another element to consider is the filmmaker's use of living subjects, and responsibility for portraying them accurately **(**Robert, 1974**)**. Without the participants the film would never succeed in reaching its goal. Therefore we must regard the participants as partners or co-authors. Profits and honours received should be shared fairly. These ideas must be always at the forefront of the filmmakers mind so to put into practice the highest sense of professional morality.

4. Conclusion

If nothing else, one should constantly comply with one's understanding of moral obligation and put the understanding into practice. The key to confronting moral challenges lies within the filmmakers' understanding of their obligations. These obligations are owed to the subjects of the documentary film and also to the wider audience. The filmmakers have to be keenly aware of this fact. Many filmmakers, however, fail to grasp the value of acting ethically and do not confront issues of obligations before plunging into the gritty world that is to be filmed. It is for this reason that this essay lays out the moral principles to be understood and the practical

steps that can be taken by the filmmakers in order to achieve the highest level of professional morality.

References

Larry Gross, John Katz, Jay Ruby (eds.). *Image Ethics: The Moral Rights of Subjects in Photographs, Film, and Television*. New York: Oxford University Press. 1998, pp. 7-8.

James Donald and Michael Renov (eds.). *The SAGE Handbook of Film Studies*. London: Thousand Oaks, Calif. 2008, pp. 147-160.

Michael Renov (eds.). *Theorizing Documentary*. New York ; London: Routledge. 2008, pp. 147-160.

Erik Barnouw. *Documentary: A History of The Non-Fiction Film*. New York: Oxford University Press. 1974, pp. 344-345.

Robert Edmonds. *About Documentary: Anthropology on Film: A Philosophy of People and Art*. Dayton, Ohio: Pflaum Pub. 1974, pp. 86-87.