

# K R I T I K E

An Online Journal of Philosophy

VOLUME 14, NUMBER 2  
DECEMBER 2020

ISSN 1908-7330

# K R I T I K E

An Online Journal of Philosophy

Volume 14, Number 2

December 2020

ISSN 1908-7330



**THE DEPARTMENT OF PHILOSOPHY**  
**University of Santo Tomas**  
**Philippine Commission on Higher Education**

---

## COPYRIGHTS

---



**All materials published by KRITIKE are licensed under a Creative Commons Attribution-NonCommercial 4.0 International License**

**KRITIKE** supports the Open Access Movement. The copyright of an article published by the journal remains with its author. The author may republish his/her work upon the condition that KRITIKE is acknowledged as the original publisher.

**KRITIKE** and the Department of Philosophy of the University of Santo Tomas do not necessarily endorse the views expressed in the articles published.

© 2007-2020 KRITIKE: An Online Journal of Philosophy | ISSN 1908-7330 | OCLC 502390973 | [kritike.editor@gmail.com](mailto:kritike.editor@gmail.com)

---

## ABOUT THE COVER

---

*KRITIKE: An Online Journal of Philosophy*, 14:2 (December 2020)

Paolo A. Bolaños, *Horizon*, 2020. Photograph.

# About the Journal

---

**KRITIKE** is the official open access (OA) journal of the Department of Philosophy of the University of Santo Tomas (UST), Manila, Philippines. It is a Filipino peer-reviewed, interdisciplinary, and international journal of philosophy founded by a group of UST alumni. The journal seeks to publish articles and book reviews by local and international authors across the whole range of philosophical topics, but with special emphasis on the following subject strands:

- **Filipino Philosophy**
- **Oriental Thought and East-West Comparative Philosophy**
- **Continental European Philosophy**
- **Anglo-American Philosophy**

The journal primarily caters to works by professional philosophers and graduate students of philosophy, but welcomes contributions from other fields (literature, cultural studies, gender studies, political science, sociology, history, anthropology, economics, inter alia) with strong philosophical content.

The word "kritike" is Greek from the verb "krinein," which means to discern. Hence, kritike means the art of discerning or the art of critical analysis. Any form of philosophizing is, in one way or another, a "critique" of something. Being critical, therefore, is an attitude common to all philosophical traditions. Indeed, the meaning of philosophy is critique and to be philosophical is to be critical.

**KRITIKE** supports the **Open Access Movement** and is classified under the "**Platinum OA**" category, which means that articles published by the journal are fully accessible online without a subscription fee. Moreover, the journal does not levy charges against the authors for the publication of their works. Articles can either be read on site or downloaded as pdf files and old issues are archived for future retrieval.

**KRITIKE** is committed to meet the highest ethical standards in research and academic publication. The journal is guided by the principles set in its **Publication Ethics and Publication Malpractice Statement**.

**KRITIKE is a Philippine Commission on Higher Education (CHED) Journal Incentive Program Recognized Journal (Journal Challenge Category)** and is indexed and listed in the following:

[The Philosopher's Index](#)  
[Web of Science™ Core Collection \(ISI\)](#)  
[Scopus® \(Elsevier\)](#)  
[Humanities International Complete™ \(EBSCO\)](#)  
[Humanities International Index™ \(EBSCO\)](#)  
[ASEAN Citation Index](#)  
[International Directory of Philosophy](#)  
[Modern Language Association \(MLA\) Directory of Periodicals](#)  
[Directory of Open Access Journals](#)  
[PhilPapers: Philosophical Research Online](#)  
[Google Scholar](#)

**KRITIKE** is a biannual journal published in June and December of each year.



# Editorial Board

---

## Editors-in-Chief

Paolo A. Bolaños, *University of Santo Tomas*  
Roland Theuas DS. Pada, *University of Santo Tomas*

## Managing Editors

Ranier Carlo V. Abengaña, *University of Santo Tomas*  
Jovito V. Cariño, *University of Santo Tomas*  
Raniel SM. Reyes, *University of Santo Tomas*

## Associate Editors

Fleurdeliz R. Altez-Albela, *University of Santo Tomas*  
Moses Aaron T. Angeles, *San Beda College*  
Marella Ada M. Bolaños, *University of Santo Tomas*  
Brian Lightbody, *Brock University*  
Peter Emmanuel A. Mara, *University of Santo Tomas*  
Melanie P. Mejia, *University of Santo Tomas*  
Dean Edward A. Mejos, *University of Asia & the Pacific*  
Mario Wenning, *University of Macau*

## Editorial Collaborative

Agustin Martin Rodriguez, *Ateneo de Manila University*  
Napoleon Mabaquiao, *De La Salle University*  
Jeffrey Oca, *Silliman University*  
Renante Pilapil, *Ateneo de Davao University*  
Ryan Urbano, *University of San Carlos*

## Book Review Editors

Darlene Demandante, *University of Santo Tomas*  
Tracy Ann P. Llanera, *Macquarie University*  
Wendyl Luna, *University of New South Wales*  
Jonathan Villacorta, *University of Santo Tomas*

## Style and Copy Editors

Gian Carla D. Agbisit, *University of Santo Tomas*  
Julia Carmela D. De Castro, *University of Santo Tomas*  
Pia Patricia P. Tenedero, *University of Santo Tomas*

## International Advisory Board

†Romualdo E. Abulad, *University of Santo Tomas*  
Karin Bauer, *McGill University*  
Bob Brecher, *University of Brighton*  
Roger Burggraeve, *Katholieke Universiteit Leuven*  
Alfredo P. Co, *University of Santo Tomas*  
William Franke, *Vanderbilt University*  
Leovino Ma. Garcia, *Ateneo de Manila University*  
Heinrich Geiger, *Katholischer Akademischer Ausländer-Dienst*  
Morny M. Joy, *University of Calgary*  
John F. X. Knasas, *University of St. Thomas – Houston*  
Megan Jane Lavery, *Columbia University*  
Zosimo E. Lee, *University of the Philippines - Diliman*  
Julius D. Mendoza, *University of the Philippines - Baguio*  
Hans-Georg Moeller, *University of Macau*  
Paul Patton, *University of New South Wales*  
Karl-Heinz Pohl, *Universität Trier*  
Peter L P Simpson, *City University of New York*  
Nicholas H. Smith, *Macquarie University*  
John Rundell, *University of Melbourne*  
†Vincent Shen, *University of Toronto*  
John Weckert, *Charles Sturt University*

# K R I T I K E

An Online Journal of Philosophy

Volume 14, Number 2

December 2020

## **A TRIBUTE TO LEONARDO N. MERCADO (1935-2020)**

1 EMMANUEL D. BATOON

A Tribute to Leonardo N. Mercado, SVD: His Legacy to the Filipino Nation

## **DEBATING DEMOCRACY**

6 JEFFRY V. OCAY

Tyranny of the Majority: Hegel and the Paradox of Democracy

19 PAOLO A. BOLAÑOS

Critical Theory and the Prospects of Radical Democracy

## **ARTICLES**

37 BOB BRECHER

Moral Cognitivism: 'Motivation' and Agency

54 MARK ANTHONY L. DACELA

Where Epistemic Safety Fails

76 IAN ANTHONY B. DAVATOS

Towards an Experimental Turn in Filipino Philosophy: A New Way Forward

97 DELFO C. CANCERAN

Saint Paul: A Revolutionary Subject

## **BOOK REVIEW**

109 JOVITO V. CARIÑO

Midgley, Mary, *What Is Philosophy For?*



————— *In Memory of* —————  
**LEONARDO N. MERCADO**  
————— 1935-2020 —————

A Tribute to Leonardo Mercado

## A Tribute to Leonardo N. Mercado, SVD: His Legacy to the Filipino Nation

---

**Emmanuel D. Batoon**

**Abstract:** This essay highlights Leonardo N. Mercado's legacy to the Filipino nation by reading his text, *Elements of Filipino Philosophy*, through the lens of critical theory. The article begins with a description of critical theory as a methodological framework in reading a text, followed by a discussion of Mercado's political interest behind his text. The essay ends by examining the relevance of his text's meaning in addressing a social crisis like the COVID-19 pandemic.

**Keywords:** Mercado, Filipino philosophy, Filipino nation, COVID-19

### Introduction

This essay highlights the legacy of Rev. Fr. Leonardo N. Mercado, SVD, Ph.D. to the Filipino nation by reading his text, *Elements of Filipino Philosophy*,<sup>1</sup> using the lens of critical theory as a methodological framework.<sup>2</sup> Critical theory is an interpretive framework that presumes that a text, such as Mercado's, does not just carry meaning but also an interest in power. Thus, to learn the interest in power and meaning of a text, the text must be situated within its historical context, where the question asked is, "Against whom and for whom is the text?"

### Against Whom is Mercado's Text?

Mercado's text was against the forces vying for power and dominance during the 1970s: on the right were the oligarchs who wanted to maintain an elitist democracy; on the left were the communists who wanted to change Philippine society through a bloody revolution conveyed in Jose Maria Sison's (under the pseudonym Amado Guerrero) book, *Philippine*

---

<sup>1</sup> Leonardo N. Mercado, *Elements of Filipino Philosophy* (Tacloban City: Divine Word University Publications, Inc., 1973).

<sup>2</sup> Emmanuel D. Batoon, *A Guide to Thesis Writing in Philosophy – Part One: Proposal Writing* (Manila: REJN Publishing, 2005), 62.



## 2 MERCADO'S LEGACY TO THE FILIPINO NATION

*Society and Revolution*;<sup>3</sup> at the periphery was the Moro National Liberation Front (MNLF) headed by Nur Misuari, who wanted to establish the Bangsamoro as an independent nation from the Filipino nation; and at the center was Ferdinand E. Marcos's government who claimed to effect a revolution from the center to build a new society through the use of Martial Law.<sup>4</sup>

Through his missionary assignments in the Philippines as an SVD (*Societas Verbi Divini* or Society of the Divine Word), Mercado encountered ethnic groups and cultures distinct from his own Visayan-Cebuano culture. Through these encounters, he noticed that people carried prejudices against other groups of people. For example, some Luzon people were biased against Visayan people, as shown in their use of Visayan people as scapegoats for whatever was wrong with their social life. Ultimately, his text was at odds with those against Filipinos for being Filipinos and who believe that "to be is to be the colonizer."<sup>5</sup>

Mercado's text resisted, as well, those who did not entertain the possibility of a philosophical alternative to the existing philosophies taught in academe during his time—Scholasticism/Thomism and existentialism. The text also opposed those who maintained academic disciplinary rigidities and territorial limits in terms of research methods.

Finally, Mercado's text opposed those who claimed that Filipino philosophy does not exist because it is not written. The text was also opposed to those who think that Filipinos do not have a sense of society because they create factions like *kami-kami* and *kayo-kayo*; that Filipinos are emotional and cannot engage in logical thinking; that they are less human than other cultures because they do not emphasize their rationality; that Filipinos do not have a moral sense; that their relationship with nature is that of domination and control; and that Filipino Catholicism is not in conformity with Orthodoxy and was not *Roman* enough.

### For Whom is Mercado's Text?

Mercado's use of the 1973 Constitution's definition of the Filipino<sup>6</sup> showed that his text was meant for Filipinos in general. In particular, he dedicated his work to the Filipino masses, as expressed in his preface to the *Elements of Filipino Philosophy*.<sup>7</sup> Thus, his text was for the oligarchs as Filipinos,

---

<sup>3</sup> Amado Guerrero, *Philippine Society and Revolution* (Hong Kong: Ta Kung Pao, 1971).

<sup>4</sup> Ferdinand E. Marcos, *Revolution from the Center: How the Philippines is Using Marital Law to Build a New Society* (Hong Kong: Raya, 1978).

<sup>5</sup> Paulo Freire, *Pedagogy of the Oppressed* (New York: Seabury Press, 1970).

<sup>6</sup> Mercado, *Elements of Filipino Philosophy*, 5.

<sup>7</sup> Mercado, Preface to *Elements of Filipino Philosophy*, xi.

excluding their marginalization of the Filipino masses. Mercado's text also welcomed the communists concerned with the Filipino masses, but never their advocacy of a bloody revolution. His text recognized the MNLF's fight for their autonomy, but not the letting go of their inclusion in the Filipino nation. Mercado's text was also a reinforcement of Ferdinand Marcos's move to establish a new society without using Martial Law as means. Thus, Mercado's text was an invitation to all political forces to focus on their being Filipinos as the common ground in building a new society.

His text was for those who see many similarities among the different Philippine ethnic groups, which can form a Filipino philosophy. The text showed this by identifying the commonalities among the Visayan, Tagalog, and Ilocano languages and behaviors. Additionally, the text indicated that the Visayans were not at all different from the rest of the Philippine ethnic groups. Ultimately, the *Elements of Filipino Philosophy* showed no other way for Filipinos "to be," but "to be" Filipinos.

Mercado's text favored those looking for an alternative philosophy to the formal academic philosophies of Scholasticism/Thomism and existentialism: Filipino cultural philosophy.<sup>8</sup> And the text justified the use of social science methods<sup>9</sup> as the appropriate method to recover the Filipino people's philosophy.

And what did the contents of the text favor?

The text favored the idea that Filipino philosophy exists. The text showed that Filipinos have a sense of society, the *sakop*, which they understand as a community of persons; the Filipinos' mind focuses on intuitive thinking because it provides them with an insight into the self or *loob* of an individual, which is the condition of possibility for their interpersonal relations; Filipinos have morality with mercy (*awa*) and concern (*malasakit*) as norms; and Filipinos extend their interpersonal relationship to nature and living in harmony with nature. Finally, Filipinos have acculturated or inculturated<sup>10</sup> Catholicism to make it their own, and Fr. Mercado elaborated this concept in his subsequent publications on Filipino theology<sup>11</sup> and Filipino religious psychology.<sup>12</sup>

---

<sup>8</sup> Emmanuel Batoon, "Tracing Mercado's Anthropological Perspective on Filipino Philosophy," in *Kritike: An Online Journal of Philosophy*, 8:1 (2014), 3.

<sup>9</sup> Leonardo N. Mercado, *Research Methods in the Philippine Context* (Tacloban City: Divine Word University Publications, Inc., 1983).

<sup>10</sup> Leonardo N. Mercado, *Inculturation and Filipino Theology* (Tacloban City: Divine Word Publications, Inc., 192).

<sup>11</sup> Leonardo N. Mercado, *Elements of Filipino Theology* (Tacloban City: Divine Word University Publications, Inc., 1975).

<sup>12</sup> Leonardo N. Mercado, *Filipino Religious Psychology* (Tacloban City: Divine Word University Publications, Inc., 1977).

#### 4 MERCADO'S LEGACY TO THE FILIPINO NATION

But are the *Elements* still relevant in explaining Filipinos' behavior during the COVID-19 pandemic?

##### **The Relevance of Mercado's *Elements of Filipino Philosophy* in Explaining Filipinos' Behavior under the COVID-19 Pandemic**

I recently attended a web-conference lecture, titled *Political Collectivism under COVID-19*, by Marc Wenyi Lai, a Taiwanese political scientist.<sup>13</sup> Lai started his lecture by showing the following statistics on COVID-19 cases and deaths worldwide as of 5 December 2020: the US had 13,000,000 cases and 200,000 deaths; India had 9,000,000 cases and 138,000 cases; Brazil had 6,000,000 cases and 173,000 deaths; Russia had 2,500,000 cases and 41,600 deaths; France had 2,000,000 and 53,000 deaths; and Spain had 1,500,000 cases and 47,000 deaths. In contrast, China had 93,797 cases and 4,000 deaths, and Taiwan had 6,861 and 81 deaths.

Eventually, Lai explained that China and Taiwan's figures were relatively low compared to the other countries because the Chinese mainlanders and the Taiwanese have a Confucian culture of obedience. This culture extends to other nations like Singapore, Vietnam, Korea, and Japan—all of which showed low COVID-19 cases and deaths. The ultimate result is that these nations possess a sense of political collectivism.

If we compare the Philippine statistics on COVID-19 cases and deaths on the same date, there were 44,000 cases and 8,572 deaths. The figures show that the Philippines is relatively nearer to China and Taiwan's figures than the US and the other nations mentioned earlier with high figures on cases and deaths. How can the statistics be explained?

Filipino philosophy can explain the Philippine statistics. The pandemic awakened the Filipinos' sense of society as *sakop*, which the authorities supported through the use of the slogans, "Pilipino kami; COVID ka lang," and "Together, we heal as one." Filipinos' intuitive thinking was reinforced, urging them to look into their fellow Filipinos' (*kababayan*) interiority (*kalooban*) and show mercy (*awa*) and concern (*malasakit*) for one another. The Filipino health workers abroad whom the British and Americans acclaimed for their care and sacrifices exemplified this moral practice. The pandemic also made Filipinos acknowledge nature's power over them and reminded them to respect and live in harmony with nature, not dominate and control nature. Finally, their sense of the sacred gave them refuge in this time of crisis.

---

<sup>13</sup> Marc Wenyi Lai, "Political Collectivism under COVID-19" (lecture presented at a web-conference organized by The Graduate School of the University of Santo Tomas, 5 December 2020).

Now that its political interest and meaning had been exposed, and its relevance in explaining the behavior of Filipinos during the COVID-19 pandemic, Mercado's *Elements of Filipino Philosophy* can be considered as a continuation of his fellow Filipino Thomasians' attempt to imagine the nation: starting from Jose Rizal, who first conceived of the Filipino nation through his novels *Noli Me Tangere* (*Touch Me Not*, 1887) and *El Filibusterismo* (*Filibusterism*, 1891), to Apolinario Mabini's "Programa constitucional de la República Filipina" ("The Constitutional Program of the Philippine Republic," 1898), down to Manuel L. Quezon who encouraged Filipinos to resist the Japanese invaders through his speech, "A Message to My People." Ultimately, Mercado's *Elements of Filipino Philosophy* stands as a testament to a life well-lived.

*Department of Political Science, University of Santo Tomas, Philippines*

## References

- Batoon, Emmanuel D., *A Guide to Thesis Writing in Philosophy – Part One: Proposal Writing* (Manila: REJN Publishing, 2005).
- \_\_\_\_\_, "Tracing Mercado's Elements of Filipino Philosophy (First of Two Parts)," in *Kritike*, 8:1 (2014).
- De Leon, Emmanuel and Marvin Einstein Mejaro, "An Interview with Leonardo Nieva Mercado, SVD," in *Kritike*, 10:2 (2016).
- Freire, Paulo, *Pedagogy of the Oppressed* (New York: Seabury Press, 1970).
- Guerrero, Amado, *Philippine Society and Revolution* (Hong Kong: Ta Kung Pao, 1971).
- Lai, Marc Wenyi, "Political Collectivism under COVID-19" (lecture presented at a web-conference organized by The Graduate School of the University of Santo Tomas, 5 December 2020).
- Marcos, Ferdinand E., *Revolution from the Center: How the Philippines Uses Martial Law to Build a New Society* (Hoong Kong: Rya, 1978).
- Mercado, Leonardo N., *Elements of Filipino Philosophy* (Tacloban City: Divine Word University Publications, Inc., 1973).
- \_\_\_\_\_, *Elements of Filipino Theology* (Tacloban City: Divine World University Publications, Inc., 1975).
- \_\_\_\_\_, *Filipino Religious Psychology* (Tacloban City: Divine Word University Publications, Inc., 1977).
- \_\_\_\_\_, *Research Methods in the Philippine Context* (Tacloban City: Divine Word University Publications, Inc., 1983).
- \_\_\_\_\_, *Inculturation and Filipino Theology* (Manila: Divine Word University Publications, Inc., 1992).



## Tyranny of the Majority: Hegel on the Paradox of Democracy

---

*Jeffry V. Ocay*

**Abstract:** At the core of the principle of democracy is the claim that all individuals, or as many as possible, should decide for themselves and that they must be included in collective governance of the community in which the majority rules. However, drawing upon Hegel's theory of the state, I will show in this paper that in a democracy, the emphasis on individual rights, at the expense of developing the notion of universal good, is not only problematic, but dangerous because in the absence of rational authority of the state, people rely mainly on public opinion for guidance, which results in what Hegel may call the tyranny of the majority. As a consequence, democracy, which purports itself to be the champion of freedom, tends to be exclusivist and totalitarian as dissenting ideas are silenced by the "ruling majority" in actual democratic processes. In fact, the notion of "legitimacy" (i.e., legitimated by the majority) conduces to the assault on the inner will to resist rendering individuals in a democracy as "conformists." The paper concludes that, for Hegel, freedom can be realized not through democracy as espoused by the liberal theorists, but through his theory of the state—the state being not only a guarantor of basic rights and liberties, but as a dimension of freedom which commits itself to a substantive vision of the universal good as the paramount object of human aspiration.

**Keywords:** Hegel, freedom, democracy, ethical life

### Introduction

The liberal theorists from Hobbes to Locke, down to Kant, Fichte, and, just recently, Rawls viewed democracy as the most fertile ground for the realization of freedom. As is well-known, the liberal theorists argued that freedom is given and that it finds expression in a democratic society. Thus, for the liberal theorists, democracy safeguards individual civil liberties through the constitution which is above the state. As we can see, this eventually makes the court the highest authority in a liberal society.

© 2020 Jeffry V. Ocay

[https://www.kritike.org/journal/issue\\_27/ocay\\_december2020.pdf](https://www.kritike.org/journal/issue_27/ocay_december2020.pdf)

ISSN 1908-7330



However, this paper shows that the contention of the liberal theorists is insufficient to account for real freedom.<sup>1</sup> I will show that instead, democracy has resulted in what Hegel may call the tyranny of the majority. I will also show that freedom is a product of a long and arduous historical process and that it finds expression not in democracy but in Hegel's theory of the state—the state being not only a guarantor of basic rights and liberties, but as a dimension of freedom which commits itself to a substantive vision of the universal good as the paramount object of human aspiration.

In what follows, I will briefly sketch Hegel's concept of freedom through a critical engagement with his seminal work, *Philosophy of Right* and then, I will present the reason why Hegel thought that liberal democracy is unsuccessful in accounting for the possibility of true freedom. In doing so, however, I will try as much as possible to avoid running the risk of being apologetic toward Hegel. Following Axel Honneth, I believe that to simply rehash the intention and argument of Hegel's *Philosophy of Right* would be anachronistic given that the social realities and style of philosophizing during Hegel's time had undergone significant changes.<sup>2</sup> For example, in the face of rigid bureaucracy, which characterizes modern societies, it would be politically naïve, if not ridiculous, to push for a return to constitutional monarchies of the early nineteenth century. Furthermore, as Honneth argues, we can no longer share Hegel's optimism that modern societies follow a continuous path of rational development since the historical development of modern societies, especially those that we witnessed during the second half of the twentieth century, had undergone significant regressions.<sup>3</sup> And lastly,

---

<sup>1</sup> Of course, there is no absolute definition of democracy. The term democracy is elastic and continues to evolve with time. The different conceptions or types of democracy also add to our difficulty in defining the term. For example, in recent years, Jürgen Habermas introduces three different normative grounds for democracy, namely, liberal, republican, and deliberative. See Odin Lysaker, "Institutional Agonism: Axel Honneth's Radical Democracy," *Critical Horizons*, 18:1, (February 2017), 33–51. Axel Honneth, another German philosopher of the Frankfurt School tradition, also offers what some scholars call a "radical conception of democracy." In his famous *Freedom's Right*, Honneth provides an interpretation of Hegel's *Philosophy of Right* that presents the ethical life as precisely the realization of the democratic ideal. See Axel Honneth, *Freedom's Right: The Social Foundations of Democratic Life*, (New York: Columbia University Press, 2014). However, its etymological definition and what it implies remain constant, that is, democracy comes from the two Greek words *demos*, which means "whole citizen living within a particular city-state, and *kratos*, which means "power or rule." Hence, etymologically speaking, democracy means "rule by the people," which necessarily implies the championing of the development and well-being of the citizens, including the protection of their rights and liberties. As I see it, Hegel is criticizing this entire conception of liberal democracy and not just a specific version of democracy which has the tendency to develop a certain form of individualistic authoritarianism.

<sup>2</sup> Axel Honneth, *Freedom's Right: The Social Foundations of Democratic Life*, trans. Joseph Ganahl (New York: Columbia University Press, 2014), 2.

<sup>3</sup> *Ibid.*, 2–3.

## 8 HEGEL ON THE PARADOX OF DEMOCRACY

the theoretical premises of philosophical discussion, according to Honneth, have also undergone a major shift since Hegel's time. In fact, Honneth stresses that the individuals in a materially enlightened era could hardly "hold onto the idealistic monism in which Hegel anchored his dialectical concept of Spirit."<sup>4</sup> Given the following caveat, my brief presentation of Hegel's concept of freedom and his critique of democracy would be purely expository.

### **Hegel's Take on Democracy: A Brief Sketch**

Perhaps the best way to articulate Hegel's opposition to democracy is to start our discussion with the nature and dynamics of democracy with emphasis on its internal contradictions. As is well-known, at the core of the principle of democracy is the claim that all individuals, or as many as possible, should decide for themselves and that they must be included in collective governance of the community in which the majority rules.<sup>5</sup> This principle presumes that political power comes from the people and that government is legitimate only when the people consent. Hence, the proponents of modern democracy believed that the creation of governments, as well as the organization of society, was made possible through a social contract. With this came the idea that people are essentially equal under natural law and that political power was derived from the people. What this implies is that for the social contract theorists, every individual possesses certain natural or inalienable rights and, therefore, is essentially free. For this reason, the social contract theorists believed that "individuals are not indebted to government or political society for their rights; rather government has its origins in the rational desires of individuals to protect their preexisting rights as human beings."<sup>6</sup> The conservative Hobbes, for example, argued that a Leviathan must be instituted to safeguard the individuals' rights and interests. Similarly, the liberal Locke, who was so particular about private property, argued that society exists for the protection of the rights of individuals (especially the right to private property).

While it can be said that democracy aims for the maximization of freedom or the expansion of autonomy, its emphasis on individual rights has rendered "human beings cut off from the external world of social settings and

---

<sup>4</sup> *Ibid.*, 3.

<sup>5</sup> For more discussion on Hegel's take on democracy, see Lucio Cortella, *The Ethics of Democracy: A Contemporary Reading of Hegel's Philosophy of Right*, trans. by Giacomo Donis (New York: SUNY Press, 2015) and Mark Tunick, "Hegel's claim about Democracy and his Philosophy of History," in *Hegel and History* (New York: SUNY Press, 2009).

<sup>6</sup> Steven B. Smith, *Hegel's Critique of Liberalism: Rights in Context*, (Chicago and London: The University of Chicago Press, 1989), 61.

institutions.”<sup>7</sup> With the absence of such social settings and institutions, which, as I will show later, is the ground for the actualization of true freedom and its eventual expansion, people rely mainly on public opinion for guidance that results in, again, what Hegel may call the tyranny of the majority. Democracy, therefore, as Hegel sees it, tends to be exclusivist as dissenting opinions are silenced in actual democratic processes.<sup>8</sup> In fact, the notion of “legitimacy” (i.e., legitimated by the majority) conduces to the assault on the inner will to resist, rendering individuals in democratic societies as “conformists.” W. G. Stratton rightly observes that the “democratic model allows for an ongoing shaping of law and social structure according to the dictate of popular sentiment.”<sup>9</sup> In this light, it is therefore not difficult for us to see why Hegel argues that democracy leads to tyranny of the majority, which may only block the actualization of true freedom. It is precisely in this respect that democracy, viewed from a Hegelian lens, is paradoxical. But what is true freedom for Hegel and how can it be actualized?

### Hegel’s Notion of the State and the Actualization of Freedom

Although individual freedom is crucial to the realization of social justice, for Hegel, it remains insufficient to account for true freedom. According to Honneth, Hegel argues that individuals need to elevate personal freedom to the level of the social, that is, individuals must also find their freedom or “self-actualization” through shared projects.<sup>10</sup> Here, Hegel’s notion of mutual recognition takes center stage. According to Honneth, mutual recognition is the key to understanding Hegel’s concept of freedom. It is worthwhile to rehash at this point Hegel’s concept of mutual recognition that he fully articulated in his seminal work, *Phenomenology of Spirit*.

In the process of the actualization of freedom, Hegel says that self-consciousness begins with “desire,” which is twofold, namely: the desire for real objects; and self-consciousness’s desire to realize itself through the realization of desire.<sup>11</sup> First of all, desire is to be understood in the

---

<sup>7</sup> Honneth, *Freedom’s Right*, 44.

<sup>8</sup> In fact, Slavoj Žižek, appropriates Hegel’s dialectic in his attack against liberal democracy. Following Hegel, Žižek argues that “exclusivity” is internal to liberal democracy. See Slavoj Žižek, “The Violence of Liberal Democracy,” *Assemblage*, 20 (April 1993), 92. See also Slavoj Žižek, *Less Than Nothing: Hegel and the Shadow of Dialectical Materialism* (London: Verso, 2012).

<sup>9</sup> W. G. Stratton, “The Problem of Democracy in Hegel’s Philosophy of Law,” *Archives for Philosophy of Law and Social Philosophy*, 74:1 (1998), 40.

<sup>10</sup> Axel Honneth, *The Pathologies of Individual Freedom: Hegel’s Social Theory*, trans. Ladislaus Lob, (Princeton and Oxford: Princeton University Press, 2010).

<sup>11</sup> It can be loosely understood as the desire of the individual to be free. Thus, as we can see in Hegel’s discussion of the master-slave dialectic, the slave is said to have the desire to be free from the master and enjoy the fruit of his own labor. Yet, because he is attached to thinghood, to things that the slave also desires, the master continues to dominate the thing by



psychological sense, for example, as a craving for something that satisfies physiological needs. But this satisfaction of need also entails the attempt of self-consciousness to assert itself as self-consciousness. Thus, desire, for Hegel, means the original attitude of the “I” as self-consciousness toward the world. In other words, desire is the necessary tendency of the acting “I” to make itself actual, that is, as free being; it is indeed the necessary self-showing of the acting “I.” The satisfaction of this desire is precisely the fulfillment of the actual Being of the “I.” However, the satisfaction of desire cannot provide the attestation of the free status that it is seeking. This is where recognition is needed. For Hegel, self-consciousness exists in itself and for itself only by being recognized by the other conscious self.<sup>12</sup> According to Herbert Marcuse, this process is a “we-like” process of Life. Here, there is an essential reciprocal dependence; in other words, there is an essential demand for reciprocal recognition.<sup>13</sup> As Robert R. Williams has shown in detail in his major studies on recognition, the “We” is a universal consciousness which results from mutual recognition, that is, when the “I” is recognized by the other “I.”<sup>14</sup> Translated into concrete social relations, mutual recognition for Hegel “refers to the reciprocal experience of seeing ourselves confirmed in the desires and aims of the other, because the other’s existence represents a condition for fulfilling our own desires and aims.”<sup>15</sup> As we can see, it is through mutual recognition, therefore, that individual freedom expands into intersubjective freedom, and expands, once again, into a “social” concept of freedom. It is important to remember, however, that Hegel’s notion of mutual recognition does not simply mean that individuals recognize themselves as free beings. Aside from the fact that before the “we” can emerge as a full

---

dominating the slave. See Donald Phillip Verene, *Hegel’s Absolute: An Introduction to Reading the Phenomenology of Spirit*, (New York: State University of New York Press, 2007), 58–59. See also Jeffrey O’Cay, “Hegel Reframed: Marcuse on the Dialectic of Social Transformation,” in *Philosophia: International Journal of Philosophy*, 16:1, (January 2015), 102–109; Jeffrey O’Cay, “Heidegger, Hegel, Marx: Marcuse and the Theory of Historicity,” *Kritike: An Online Journal of Philosophy*, 2:2 (December 2008), 46–64; Jeffrey O’Cay, “Ethics of Refusal: Globalization and the Penan People’s Struggle for Recognition,” in *Budhi: A Journal of Ideas and Culture*, 19:2&3, (2015), 169–195; and Jeffrey O’Cay, “Eroticizing Marx, Revolutionizing Freud: Marcuse’s Psychoanalytic Turn”, in *KRITIKE: An Online Journal of Philosophy*, 3:1, (June 2009), 10–23.

<sup>12</sup> *Ibid.*, 229.

<sup>13</sup> Herbert Marcuse, *Hegel’s Ontology and the Theory of Historicity*, translated by Seyla Benhabib, (Massachusetts: The MIT Press, 1987), 251.

<sup>14</sup> Robert R. Williams, Translator’s Introduction to Georg Wilhelm Friedrich Hegel, *Lectures on the Philosophy of Spirit, 1827–8*, trans. by Robert R. Williams (New York: Oxford University Press, 2007), 19. See also Robert R. Williams, *Recognition: Fichte and Hegel on the Other* (Albany: SUNY Press, 1992).

<sup>15</sup> Honneth, *Freedom’s Right*, 45. See also Jean-Philippe Deranty, *Beyond Communication: A Critical Study of Axel Honneth’s Social Philosophy* (Leiden: E.J. Brill, 2009), and Jean-Philippe Deranty and Emmanuel Renault, “Politicizing Honneth’s Ethics of Recognition,” *Thesis Eleven*, 88:1 (February 2007), 92–111.

community of recognition, a specific dialectic must be gone through which involves the famous concept of the struggle for recognition. For Hegel, mutual recognition requires a specific social and cultural context upon which such recognition is grounded. In other words, mutual recognition needs to be mediated by institutions and social practices, which Hegel calls “institutions of ethical life” or *Sittlichkeit*. Hegel claims that it is only through the institutions of ethical life, in which the individuals participate, can they recognize each other as free beings.<sup>16</sup> This is because the primary role of the institutions of ethical life is to preserve and enhance the individuals’ right to mutual recognition.

The ethical life for Hegel, therefore, is the consummate actualization of freedom. As Hegel writes in the concluding sentence of Paragraph 142 of the *Philosophy of Right*: “The ethical system is thus the conception of freedom developed into a present world, and also into the nature of self-consciousness.”<sup>17</sup> And in Paragraph 143, Hegel further writes:

The conception of the *universal* will, when united with the realization of the will, or the particular will, is knowing. Hence arises the consciousness of the distinction between these two phases of the idea. But the consciousness is now present in such a way that each phase is separately the totality of the idea, and has the idea as its content and foundation.<sup>18</sup>

What this passage amounts to is that Hegel’s notion of the ethical life as the consummate actualization of freedom is the reconciliation of subjective will and universal will. Here, the universal will, which Hegel understands as custom, does not appear as external to the individual will, but becomes a second nature, which takes the place of the original and merely natural will.<sup>19</sup> And for Hegel, once the universal will has taken the place of the original and merely natural will, it has become “the very soul, meaning, and reality of one’s life.”<sup>20</sup> Put differently, in the ethical life, the individual consciously obeys the laws of the society (or, in Talcott Parson’s words, internalize the customs of the society) because she is convinced that such laws are

---

<sup>16</sup> See Alan Patten, *Hegel’s Idea of Freedom* (Oxford: Oxford University Press, 2005), 125–130. See also Gavin Rae, “Realizing Freedom: Hegel and Ethical Life,” in *Realizing Freedom: Hegel, Sartre, and the Alienation of Human Being* (London: Palgrave Macmillan, 2011), 183–230.

<sup>17</sup> G. W. F. Hegel, *Philosophy of Right*, trans. by S. W. Dyde (Mineola, New York: Dover Publications, 2005), 76. The text will hereafter be cited by the page number. References to the paragraph number will be indicated in the body of the text.

<sup>18</sup> *Ibid.* Italics mine.

<sup>19</sup> *Ibid.*, 80.

<sup>20</sup> *Ibid.*

expressions of the universal will. The conformity of the particular will with the universal will, therefore, is what Hegel meant by true freedom, one that is ruled out in liberalism's individualistic conception of freedom. And it is interesting to note that Hegel calls the reconciliation of particular will and universal will "Spirit." As Hegel writes in Paragraph 151 of the *Philosophy of Right*, the ethical life in the form of custom "is the living spirit actualized as a world; by this actualization does the substance of spirit exist as spirit."<sup>21</sup> This contention provides us with the key to unlock one of the mysteries of Hegel's monumental work, *Phenomenology of Spirit*: that the development of *Geist* (Spirit) is nothing but the development of freedom actualized in the state.

### Freedom and the Three Moments of the Ethical Life

Now, if we recall, Hegel says that the ethical life, as the reconciliation of the universal will and particular will, is the idea of freedom developed into the existing world and nature of self-consciousness. This means for Hegel that the ethical life is a process of development. In Paragraph 157 of the *Philosophy of Right*, Hegel writes: "The conception of this idea (ethical life) exists only as spirit, as active self-knowledge and reality, since it objectifies itself by passing through the form of its elements."<sup>22</sup> Hegel now describes the three moments that the ethical life passes through, namely: the family, civil society, and the state.

The family, for Hegel, represents ethical life in its simplest form because the unity of the family is an immediate one based on the feeling of love. In the family, members do not relate to each other as independent but as parts of a larger whole to which they immediately identify. In Robert Pippin's interpretation of Hegel's concept of the family as the first phase of the ethical life, the family is viewed as an ethically binding institution not because it is primarily a natural institution—that is, as the basic unit of society—but because we see in the family an active recognition of mutual dependencies that is a necessary requirement for the realization of independence.<sup>23</sup> However, the family also exists in a larger context where members are not bound together by the natural feeling of love. This larger context for Hegel is the realm of civil society—the sphere of economic or market relations. Although we could hardly consider the economic system today as a realm of social freedom, given the individualistic nature of the profit-seeking capitalist system, Hegel thought that individuals in the sphere

---

<sup>21</sup> *Ibid.* See also Paul Franco, *Hegel's Philosophy of Freedom* (New Haven and London: Yale University Press, 1999), 229.

<sup>22</sup> Hegel, *Philosophy of Right*, 82. Insertion mine.

<sup>23</sup> Robert B. Pippin, "What is the Question for which Hegel's Theory of Recognition is the Answer?," *European Journal of Philosophy*, 8:2 (2000), 166.

of market relations, of course during his time, are implicitly governed by the universal which invisibly led to serving each other's interest. As Hegel writes in Paragraph 183 of the *Philosophy of Right*:

The self-seeking end is conditioned in its realization by the universal. Hence is formed a system of mutual dependence, a system which interweaves the subsistence, happiness, and rights of the individuals with the subsistence, happiness, and rights of all."<sup>24</sup>

Just as in the family, Hegel argues that the civil society as the second phase of the ethical life is, therefore, a realm of social freedom because we also see in it a kind of mutual recognition that individuals need for them to become truly free. But Hegel argues that although the civil society seems to have everything it needs for it to become complete, such as a justice system "to regulate the interactions between individuals as they pursue their economic interests, a police to maintain public order, a public authority to regulate the market and provide for the poor, and a corporations to lift individuals out of their narrow interests and lead them to identify with a more universal purpose,"<sup>25</sup> the type of freedom it harbors remains subjective. It is for this reason that Hegel moves on to the state in order to bring into the picture the actualization of true freedom, that is, as already mentioned, the reconciliation of the universal will and the particular will.

In Paragraph 257 of the *Philosophy of Right*, Hegel defines the state as the actuality of the ethical spirit.<sup>26</sup> He goes on to say that "it is the will which manifests itself, makes itself clear and visible, substantiates itself. It is the will which thinks and knows itself, and carries out what it knows, and in so far as it knows."<sup>27</sup> According to Paul Franco, what this typically abstract phrase indicates is the essential relationship between state and freedom; the former being the actuality of the latter.<sup>28</sup> In other words, for Hegel, the state is the concrete embodiment of true freedom. But freedom here, Franco notes, should not be understood as the particular or arbitrary will of the individual, but as the individual's substantial or rational will. What this means is that true freedom for Hegel is a kind of freedom that one enjoys in being with oneself in another, which consists in the unity of subjective and objective

---

<sup>24</sup> Hegel, *Philosophy of Right*, 97. See also Timothy C. Luther, *Hegel's Critique of Modernity: Reconciling Freedom and Community*, (United Kingdom: Lexington Books, 2009), 156.

<sup>25</sup> Franco, *Hegel's Philosophy of Freedom*, 280. See also Kenneth Kierans, "The Concept of Ethical Life in Hegel's *Philosophy of Right*," *History of Political Thought*, 12:3 (Autumn 1992), 417–435 and Allen W. Wood, *Hegel's Ethical Thought*, (New York: Cambridge University Press, 1990).

<sup>26</sup> Hegel, *Philosophy of Right*, 132.

<sup>27</sup> *Ibid.*

<sup>28</sup> Franco, *Hegel's Philosophy of Freedom*, 283.

freedom. On the one hand, Hegel says that an individual is subjectively free if she reflects on her actions rather than blindly acting on the authority of public opinion. On the other, Hegel says that an individual is objectively or substantially free if her actions accord with reason, that is, if she acts in accordance with the tasks and functions (e.g., voting during elections or paying taxes) that are asked of her as a good citizen of the state. With this, we can say that being able to decide for oneself as a result of reflection is not yet true freedom for Hegel as “rational self-determination.” Subjective freedom needs to pass over into objectivity, into an objective set of principles of action that individuals find themselves committed to. According to Honneth’s interpretation of Hegel’s concept of freedom, “a subject is only ‘free’ if it encounters another subject, within the framework of institutional practices, to whom it is joined in a relationship of mutual recognition.”<sup>29</sup> This is exactly what Hegel meant by the reconciliation of the particular or individual will and the universal will. For Hegel, therefore, the state as the actuality of the ethical spirit, or the actuality of concrete or true freedom, is not only a restraint on freedom, but “the necessary context within which our individual powers and capacities can grow and develop.”<sup>30</sup>

Now, it is important to remember that in Hegel’s concept of the ethical life, the individual’s obedience to laws is not a blind and unconditional obedience. The individual obeys the laws because she knows that the laws are just and, hence, rational. Thus, true freedom for Hegel also implies possessing personal knowledge of reality, for example, knowing what is right and wrong. Allan Patten says that individuals do not just believe in ethical principles simply because they are laid down by external laws and precepts of authority, but individuals should have assent, disposition, conscience, and full awareness of such principles.<sup>31</sup> In fact, according to Honneth, “Hegel did not merely wish to affirm and reinforce current practices and institutions, but also to correct and transform them.”<sup>32</sup> Here, we have an awareness of the rationality of the laws because we obey on the basis of our own insight and reason, and not on the authority of public opinion and unexamined will. This is now the problem with which I would like to close this paper.

---

<sup>29</sup> Honneth, *Freedom’s Right*, 45.

<sup>30</sup> Smith, *Hegel’s Critique of Liberalism*, xi.

<sup>31</sup> Patten, *Hegel’s Idea of Freedom*, 66.

<sup>32</sup> Honneth, *Freedom’s Right*, 8. Thus, Hegel’s famous line “What is rational is real; and what is real is rational” does not mean that we are always in the midst of the “best of all possible worlds,” that what exists is always actual and rational. Since actuality or reality for Hegel is always the unity of the universal and particular, then it (actuality) remains to be realized. It could be a mere potentiality that needs to be realized, to use Aristotelian terminology. In fact, for Hegel, a bad state is not a genuine reality. It merely exists. See Hegel, *Philosophy of Right*, 152-153. See also Trent Schroyer, *The Critique of Domination: The Origins and Development of Critical Theory* (Boston: Beacon Press, 1973), 55.

## The Paradox of Democracy

First of all, Hegel's misgiving about democracy stems from his conception of the ethical life. To reiterate, in the Hegelian conception of the state as the actuality of freedom, of the ethical spirit, we see in it a universal interest that does not sacrifice individual rights and liberties. We see in Hegel's conception of the state the actualization of Plato's notion of the Good. As already mentioned, democracy has overemphasized individual rights and liberties at the expense of developing the notion of public good. Hence, in this model of government, the state is seen simply as the guarantor of life, property, and liberty. In other words, the state in a democracy is not an "end in itself and for itself," but simply as a means for the promotion of individual interests. For this reason, democracy, as Hegel would have us believe, becomes a fertile ground for the expansion and maximization not of freedom, but of greed. With this idea in mind, it is therefore not difficult for us to see why liberal democracy goes hand in hand with capitalism.

A second reason why Hegel sees democracy as paradoxical is that "in a democracy, people are both judges and parties in the administration of law, and this poses the greatest danger to the rights of individuals and the rule of law."<sup>33</sup> Because the key intuition in democracy is the idea that the people are the basis of political legitimacy, and individual right is the supreme value, this means that democracy abandons the rational authority of the state in favor of public opinion. As a consequence, democracy, which purports itself to be the champion of freedom, tends to be exclusivist and totalitarian as dissenting ideas are silenced by the "ruling majority" in actual democratic processes. In fact, the notion of "legitimacy" (i.e., legitimated by the majority) conduces to the assault on the inner will to resist, which renders individuals in a democracy as "conformists." As the famous American thinker Allan Bloom corroborates:

The most successful tyranny is not the one that uses force to assure uniformity but the one that removes the awareness of other possibilities, that makes it seem inconceivable that other ways are viable, that removes the sense that there is an outside. It is not feelings or commitments that will render a man free, but thoughts, reasoned thoughts. Feelings are largely formed and informed by convention. Real differences come from difference in thought and fundamental principle. Much

---

<sup>33</sup> See Franco, *Hegel's Philosophy of Freedom*, 27.



in democracy conduces to the assault on awareness of difference.<sup>34</sup>

Thus, as I noted previously, democracy's emphasis on individual rights at the expense of developing the notion of universal good is dangerous because in the absence of the rational authority of the state, people will rely mainly on public opinion for guidance. As I see it, it would seem that Hegel is making the claim that it is indeed a tyranny of the majority: a situation that is characterized primarily by the failure of democratic societies to promote the realization of true freedom. In a footnote to Paragraph 317 of the *Philosophy of Right*, Hegel quoted Goethe saying: "Zuschlagen kann die Masse, Da ist sie respektabel; Urtheilen gelingt ihr miserabel."<sup>35</sup> To reiterate, democracy for Hegel is therefore paradoxical because, as we can see in the foregoing discussion, it is incapable of manifesting the concretion of true freedom.

At this point, I think it helps if we rehash the key intuition of Hegel's theory of the state to offer some alternative to the tyranny of the majority. For Hegel, the antidote to the tyranny of the majority is the rational authority of the state. Thus, in Paragraph 258 of the *Philosophy of Right*, Hegel argues that the highest duty of the individual is to be a member of the state. It might be worth quoting some of the longer extracts of the Note to this paragraph:

Were the state to be considered as exchangeable with the civic society, and were its decisive features to be regarded as the security and protection of property and personal freedom, the interest of the individual as such would be the ultimate purpose of social union. It would then be at one's opinion to be a member of the state. But the state has a totally different relation to the individual. It is the objective spirit, and he has his truth, real existence, and ethical status only in being a member of it. Union, as such, is in itself the true content and end, since the individual is intended to *lead* a universal life.<sup>36</sup>

For Hegel then, freedom can be realized not through democracy as espoused by the liberal theorists, but through his theory of the state—the state being not only a guarantor of basic rights and liberties, but as a dimension of freedom which commits itself to a substantive vision of the universal good as

---

<sup>34</sup> Allan Bloom, *The Closing of the American Mind*, (New York: Simon & Schuster, 1987), 249.

<sup>35</sup> "The masses are respectable hands at fighting, but miserable hands at judging." Goethe, "Sprichtwörtlich" (1825), as quoted in Hegel, *Philosophy of Right*, 188.

<sup>36</sup> Hegel, *Philosophy of Right*, 133. Italics mine.

the paramount object of human aspiration. It is precisely in the context of this unique Hegelian conception of freedom that many social and political theorists today have sought refuge in Hegel rather than in Kant who champions the idea of individual autonomy. As Steven Smith writes, “If contemporary liberals, such as Rawls, have been led to rediscover Kant, liberalism’s critics have been forced to reinvent Hegel.”<sup>37</sup>

*The Graduate School, Eastern Visayas State University  
Tacloban City, Philippines*

## References

- Bloom, Allan, *The Closing of the American Mind* (New York: Simon & Schuster, 1987).
- Cortella, Lucio, *The Ethics of Democracy: A Contemporary Reading of Hegel’s Philosophy of Right*, trans. by Giacomo Donis, (New York: SUNY Press, 2015).
- Deranty, Jean-Philippe, *Beyond Communication: A Critical Study of Axel Honneth’s Social Philosophy* (Leiden: E.J. Brill, 2009).
- Deranty, Jean-Philippe and Emmanuel Renault, “Politicizing Honneth’s Ethics of Recognition,” *Thesis Eleven*, 88:1 (February 2007).
- Franco, Paul, *Hegel’s Philosophy of Freedom* (New Haven and London: Yale University Press, 1999).
- Hegel, G. W. F., *Philosophy of Right*, trans. by S. W. Dyde, (Mineola, New York: Dover Publications, 2005).
- Honneth, Axel, *Freedom’s Right: The Social Foundations of Democratic Life*, trans. by Joseph Ganahl, (New York: Columbia University Press, 2014).
- \_\_\_\_\_, *The Pathologies of Individual Freedom: Hegel’s Social Theory*, trans. by Ladislaus Löb (Princeton and Oxford: Princeton University Press, 2010).
- Kierans, Kenneth, “The Concept of Ethical Life in Hegel’s *Philosophy of Right*,” in *History of Political Thought*, 12:3 (Autumn 1992).
- Luther, Timothy C., *Hegel’s Critique of Modernity: Reconciling Freedom and Community*, (United Kingdom: Lexington Books, 2009).
- Lysaker, Odin, “Institutional Agonism: Axel Honneth’s Radical Democracy,” in *Critical Horizons*, 18:1 (February 2017), 33–51.
- Marcuse, Herbert, *Hegel’s Ontology and the Theory of Historicity*, trans. by Seyla Benhabib, (Massachusetts: The MIT Press, 1987).

---

<sup>37</sup> Smith, *Hegel’s Critique of Liberalism*,” 4. Italics mine.

- Ocay, Jeffrey, "Hegel Reframed: Marcuse on the Dialectic of Social Transformation," in *Philosophia: International Journal of Philosophy*, 16:1 (January 2015).
- \_\_\_\_\_, "Heidegger, Hegel, Marx: Marcuse and the Theory of Historicity," in *Kritike: An Online Journal of Philosophy*, 2:2 (December 2008).
- \_\_\_\_\_, "Ethics of Refusal: Globalization and the Penan People's Struggle for Recognition," in *Budhi: A Journal of Ideas and Culture*, 19:2&3 (2015).
- \_\_\_\_\_, "Eroticizing Marx, Revolutionizing Freud: Marcuse's Psychoanalytic Turn," in *Kritike: An Online Journal of Philosophy*, 3:1 (June 2009).
- Patten, Allen, *Hegel's Idea of Freedom*, (Oxford: Oxford University Press, 2005).
- Pippin, Robert B., "What is the Question for which Hegel's Theory of Recognition is the Answer?," in *European Journal of Philosophy*, 8:2 (2000)..
- Rae, Gavin, "Realizing Freedom: Hegel and Ethical Life," in *Realizing Freedom: Hegel, Sartre, and the Alienation of Human Being* (London: Palgrave Macmillan, 2011), 183–230.
- Schroyer, Trent, *The Critique of Domination: The Origins and Development of Critical Theory* (Boston: Beacon Press, 1973).
- Smith, Steven B., *Hegel's Critique of Liberalism: Rights in Context* (Chicago and London: The University of Chicago Press, 1989).
- Stratton, W. G., "The Problem of Democracy in Hegel's Philosophy of Law," in *Archives for Philosophy of Law and Social Philosophy*, 74:1 (1998).
- Tunick, Mark, "Hegel's Claim about Democracy and his Philosophy of History," in *Hegel and History*, (New York: SUNY Press, 2009).
- Verene, Donald Philip, *Hegel's Absolute: An Introduction to Reading the Phenomenology of Spirit*, (New York: State University of New York Press, 2007).
- Williams, Robert R., *Recognition: Fichte and Hegel on the Other*, (Albany: SUNY Press, 1992).
- \_\_\_\_\_, Translator's Introduction to Georg Wilhelm Friedrich Hegel, *Lectures on the Philosophy of Spirit, 1827–8*, trans. by Robert R. Williams (New York: Oxford University Press, 2007).
- Wood, Allan W., *Hegel's Ethical Thought*, (New York: Cambridge University Press, 1990).
- Žižek, Slavoj, "The Violence of Liberal Democracy," *Assemblage*, Vol. 20 (April 1993), 92–93.
- \_\_\_\_\_, *Less Than Nothing: Hegel and the Shadow of Dialectical Materialism*, (London: Verso, 2012).

## Critical Theory and the Prospects of Radical Democracy

---

*Paolo A. Bolaños*

**Abstract:** In this paper I emphasize the link between Honneth's critical theory and radical democracy as defined by C. Douglas Lummis. I, firstly, present Lummis's portrayal of radical democracy, emphasizing the original meaning of the notion of democracy as essentially radical in contrast to muddled conceptions of democracy. I, then, briefly present a characterization of radical democracy as a philosophical and normative principle. I emphasize, following Lummis, that what is radical in democracy is common sense language that collectively binds people. I relate this to Hegel's idea of *Sittlichkeit*. Gesturing towards the idea that democracy is a kind of participative discourse, I propose that Honneth's theory of social freedom is a third possibility between Habermas's deliberative discourse and Mouffe's agonistic discourse. I, then, rehearse the three normative claims of Horkheimer to contextualize Honneth's commitment to critical theory, allowing me to present a schematic account of his theory of social freedom which is ironically Hegelian inspired, but decidedly critical of Hegel's characterization of democracy. I conclude by relating Benjamin's image of "the tradition of the oppressed" with the notion of social freedom.

**Keywords:** critical theory, radical democracy, social freedom, justice

### Muddled Democracy

Democracy comes from the Greek words *dêmos* ("people") and *krátos* ("power"), literally it means "people power." The Greek idea of *dêmos* referred to the "poorest and numerous class of citizens."<sup>1</sup> In the *Constitution of Athens*, Aristotle chronicles how the statesman, Solon, alleviated the crisis of debt that the poor Athenians suffered under the

---

<sup>1</sup> C. Douglas Lummis, *Radical Democracy* (Ithaca: Cornell University Press, 1996), 15.

wealthy of Athens.<sup>2</sup> Democracy, therefore, evolved in ancient Greece as a reproach to debt and debt-bondage. As such, *dêmos* referred to the oppressed class which constituted majority of the Athenian citizenry. In this context, “democracy” or “people power” meant for the Greeks freedom from economic bondage via the empowerment of the most numerous of citizens. This radically anticipates one of the basic tenets of Marxist philosophy, that is, the proletariat’s freedom from economic slavery.

From the very beginning, therefore, the idea of democracy presupposes the idea of “the people” and their right to live in a society free from oppression of any ruling class that does not represent the majority. As such, democracy “is a critique of centralized power of every sort—charismatic, bureaucratic, class, military, corporate, party, union, technocratic. By definition it is the antithesis to all such power.”<sup>3</sup> The situation here in the Philippines is quite ironic, however. Democracy, as a theory and practice, is understood based on problematic presuppositions about the idea of “the people.”

C. Douglas Lummis, in his book *Radical Democracy*, points out three instances of misunderstanding the idea of “the people” which will help me illustrate this point. Firstly, there is a misunderstanding of the idea when “the people” is construed to represent the middle and upper classes. This is an instance where democracy is invoked in order to justify the interest of the middle and upper classes, while the lower class—that is to say, the class of servants and laborers—that maintain the surplus of wealth of the former classes is not recognized as a legitimate contributor to the practice of democracy. Secondly, the idea of “the people” is misconstrued as those groups of people who support a particular political party that purports to advocate democratic ideals. Within this framework, the politics of inclusion-exclusion is at play—those who adhere to the political party are construed to be advocates of democracy, while those who do not may be seen as “enemies” of democracy, that is to say, “enemies of the people.” In such instance, the “enemies of the people” may either meet some level of aggression from

---

<sup>2</sup> Aristotle writes: “it came to pass that the upper classes and the people were divided by party-strife for a long period, for the form of government was in all respects oligarchical; indeed, the poor were in a state of bondage to the rich, both themselves, their wives, and their children, and were called *Pelatae* (bond-slave for hire), and *Hektemori* (paying a sixth of the produce as rent); for at this rate of hire they used to work the lands of the rich. Now, the whole of the land was in the hands of a few, and if the cultivators did not pay their rents, they became subject to bondage, both they and their children, and were bound to their creditors on the security of their persons, up to the time of Solon. For he was the first to come forward as the champion of the people. The hardest and bitterest thing then to the majority was that they had no share in the offices of government; not but what they were dissatisfied with everything else, for in nothing, so to say, had they any share.” Aristotle, *The Constitution of Athens*, trans. by Thomas J. Dymes (London: Seeley and Co., Limited, 1891), 1-2.

<sup>3</sup> Lummis, *Radical Democracy*, 25.

members of the political party or they are simply ignored. Thirdly, a variation of the second is when a party or the state presumes that it is the “voice” of the people. However, this often results in the reification of the concept of the people because, instead of actual people governing themselves, representational politics becomes the dominant practice.<sup>4</sup>

The muddled connotations of “the people” mentioned above inform the way Filipinos understand democracy. In the Philippines (probably also elsewhere) the so-called democratic process seems to work for the middle and upper classes. The lower class, albeit constitutes the largest number in terms of voting power, seems to be largely excluded as their votes translate into the election of politicians whose status in society is established by political lineage or by sheer celebrity status. Moreover, the second feature of muddled democracy is also observable in Philippine society. Perhaps the most palpable example is the culture of “bloc voting” by particular groups, especially the religious ones.<sup>5</sup> Within these religious groups, bloc voting is enacted as a result of the doctrine that no member of the Church should destroy the unity of the Church by voting otherwise than who the Church leaders anoint as their official political candidate. The doctrine is a double-edged sword inasmuch as, on the one hand, the influence of the Church on a government that it helped established is fortified and, on the other hand, bloc voting becomes favorable to politicians aspiring for positions.<sup>6</sup> There is no room for deliberation and criticism on the part of the Church members, only the vicarious transfer of power, lest one risks expulsion. In terms of the third feature of muddled democracy mentioned above, the political system in the Philippines allows political parties galore. However, despite this mechanism, unnecessarily allowing too many parties is self-defeating. Because every party is saying almost the same thing and advocating almost the same ideals. For instance, what we have is a repetition, through public relations, of the narrative of the plight of the poor, resulting in the reification of the narrative and trivialization of the participatory potential of the people.

At the end of day, politics in the Philippines is endemically patronage-based politics. So, as oppose to a way of life or an ideal, democracy is reduced into a “method” of determining who shall govern. In the Philippines, such method has maintained the culture of political cronyism, patronage, and dynasty.

---

<sup>4</sup> Cf. *Ibid.*, 15-16.

<sup>5</sup> Gerg Cahiles, “Politics of recognition: The power of religious endorsement in the 2019 polls,” in *CNN Philippines* (6 April 2019), <<https://cnnphilippines.com/news/2019/4/10/religious-endorsement-philippine-elections.html>>.

<sup>6</sup> Fiona Nicolas, “Duterte, Marcos get INC endorsement for May 9 polls,” in *CNN Philippines* (5 May 2016), <<https://cnnphilippines.com/news/2016/05/05/Iglesia-ni-Cristo-endorsement-duterte-marcos-May-9-elections.html>>.



### Radical Democracy as Philosophy and Normative Principle

The critique of the muddled form of democracy described above would be within the ambit of “radical democracy,” a name used by Lummis to describe what he deems as “democracy in its essential form” as opposed to “modified democracies.”<sup>7</sup> Perhaps, it is important to note here that the notion of radical democracy I am reconstructing is different from a notion of radical democracy based on the idea of sheer “political will” demonstrated by strong leadership.<sup>8</sup> Following Lummis, I wish instead to illustrate something more fundamental, that is to say, truly “radical” (from the Latin *radix* which means “root”), to the idea and practice of democracy. Lummis highlights the main features of radical democracy and refers to the idea as a going back to the original meaning of democracy and a veering away from notions of democracy prefixed by modifiers, such as, “liberal,” “social,” “Christian,” etc. It is important to note that the term “radical” is used, not to modify, but to intensify the fundamental feature of democracy — that it is, first and foremost, radical. By radical, Lummis means that democracy has been, from its very inception in ancient Greece, “a critique of centralized power of every sort ... an antithesis to such power.”<sup>9</sup> Lummis further writes:

... radical democracy is subversive everywhere. It is subversive not only in military dictatorships but also in the countries that are called democratic, those that are called socialist, and those that are “postsocialist.” It is subversive not only inside the big corporations but also inside the big unions. It is the idea that joins people struggling for liberty in all countries and all situations — if only they could all see in that way.<sup>10</sup>

Apart from a “critical attitude,” democracy, says Lummis, is informed by the language of “common sense.” This means that democracy, as a way of life and ideal, can be expressed in ordinary language. Lummis, however, points out that ordinary language is more complex than the language of intellectuals, like philosophers and social scientists. “Technical terms are supposed to refer only to specific and clearly defined meanings, whereas the words of ordinary language bear all the complexity of the disorderly history

---

<sup>7</sup> Lummis, *Radical Democracy*, 24-25.

<sup>8</sup> See, for instance, Christian Ryan Maboloc, “President Rodrigo Duterte and the Birth of Radical Democracy in the Philippines,” in *International Journal of Politics and Security*, 2:3 (May 2020), 116-134.

<sup>9</sup> Lummis, *Radical Democracy*, 24-25.

<sup>10</sup> *Ibid.*

of their uses.”<sup>11</sup> Common sense is a product of the collective use of ordinary language. Hence, as opposed to a discourse carried out only by intellectual elites in the academe, the discourse of democracy must be carried out in ordinary language.<sup>12</sup>

Structured by common sense language, a language that collectively binds people, democracy, Lummis argues, is a “moral discourse.” This entails that democratic discourse is not a random convergence of people, but, rather, it is carried out through moral discourse, choice, and action—it is a result of cooperation, joint action, and agreement, normatively based on a common cause. As such, democracy should not be reduced into a method, but democracy is rather the goal, the ideal. As a goal or ideal, it presupposes the participation of all members of a community. What kind of participation is required by democratic discourse is a question that is yet to be answered; but we can imagine, for a moment, that it could perhaps assume either the deliberative discourse described by Jürgen Habermas or the agonistic discourse described by Chantal Mouffe. For Habermas, democratic discourse is possible through the participation of different parties that accept a given set of rules of action aimed at the realization of a specific end. It is by virtue of this participative agreement that a given set of rules of action is deemed justified. Habermas declares that, “Only those norms can claim to be valid that meet (or could meet) with the approval of all affected in their capacity *as participants in a practical discourse*.”<sup>13</sup> Meanwhile, Mouffe takes the notion of democracy a step further by arguing that any type of deliberation rests on a political ontology of agonism she calls “agonistic pluralism.” For Mouffe, any identity rests on a relation of difference, that is, the affirmation of the difference of the other (constitutive outside), thereby setting the stage for antagonisms.<sup>14</sup> In this context, Mouffe understands antagonism or conflict as the very constitutive element of the democratic process and says that antagonism “is inherent in all human society” and “can take many different forms and can emerge in diverse social relations.”<sup>15</sup> Inasmuch as antagonism is inherent in all human relations, Mouffe, nevertheless, argues that the “aim of democratic politics is to transform an ‘antagonism’ into an ‘agonism’.”<sup>16</sup> In

---

<sup>11</sup> *Ibid.*, 21.

<sup>12</sup> Cf. *Ibid.*

<sup>13</sup> Jürgen Habermas, *Moral Consciousness and Communicative Action*, trans. by Christian Lenhardt and Shierry Weber Nicholsen (Cambridge: Polity Press, 1990), 66.

<sup>14</sup> Mouffe writes: “When we accept that every identity is relational and that the condition of existence of every identity is the affirmation of a difference, the determination of an ‘other’ that is going to play the role of a ‘constitutive outside’, it is possible to understand how antagonisms arise.” Chantal Mouffe, *The Return of the Political* (London: Verso, 1993), 2.

<sup>15</sup> Chantal Mouffe, “Deliberative Democracy or Agonistic Pluralism?,” in *Social Research*, 66:3 (Fall 1999), 754.

<sup>16</sup> *Ibid.*, 755.

this context, the goal of democracy, then, “is not to eliminate passions ... in order to render them rational consensus possible, but to mobilise those passions towards the promotion of democratic designs.”<sup>17</sup>

If we juxtapose the respective positions of Habermas and Mouffe with that of Lummis, then the importance of language is apparent. Clearly, the Habermasian idea of practical discourse presupposes some sort of common language shared by all participants in a given discourse, the purpose of which is the practical hope of reaching a common understanding of a given situation leading to what he calls “communicative action.”<sup>18</sup> Meanwhile, despite the fact that Mouffe is veering away from a deliberative discourse grounded in the a common rationality of the Habermasian sort, her agonistic pluralism, nevertheless, still presupposes that identities are constructed or deconstructed through the agonistic (dialectical) interaction among subjects or groups. Agonistic struggles, according to Mouffe (and Ernesto Laclau), are conditioned by *forms of articulation* that establish “a relation among elements such that their identity is modified ... The differential positions, insofar as they appear articulated within a discourse we will call *moments*.”<sup>19</sup> In this context, the agonistic interaction among moments—articulated as *fixed elements*—presupposes a relational ontology that decenters the fixed articulation of elements, as such, the agonal discourse demonstrates the impossibility of the total or final closure of an interaction. This is so because the interaction of elements is not grounded in a fixed totality, but, rather, in contingency. I could perhaps cite an example from the history of the Philippines during the past five decades. The declaration of Martial Law by Ferdinand Marcos in 1972 is an instance of an articulation of an element as a reaction to the threats of the Communist Party of the Philippines (CPP) and Mindanao Independence Movement (MIM). Marcos’s declaration came not without any “antagonisms” from oppositional figures and the reign of martial rule itself was marred by instances of human rights violations, corruption, and various abuses. These antagonisms exposed the internal limitations of martial rule which finally resulted in the ouster of Marcos during the 1986 People Power event—a moment in the history of the Philippines that “disarticulated” the tyrannical Marcos regime. In recent years, however, it seems like Filipinos have been witnessing a “rearticulation” of martial rule, a climate that creates the opportunity for critique and disarticulation; the suppression of critique is of course to be

---

<sup>17</sup> *Ibid.*, 756.

<sup>18</sup> See Jürgen Habermas, *The Theory of Communicative Action*, Vol. 1, trans. by Thomas McCarthy (Boston: Beacon Press, 1984), 273-337.

<sup>19</sup> Ernesto Laclau and Chantal Mouffe, *Hegemony & Socialist Strategy: Towards a Radical Democratic Politics* (London: Verso, 1985), 105.

expected.<sup>20</sup> For Mouffe, this sort of antagonism is a “negation of a given order,” one that “is, quite simply, the limit of that order, and not a moment of a broader totality.”<sup>21</sup>

Given the above, the difference between Habermas and Mouffe is that the former advocates a version of democratic discourse based on a shared rationality, while the former is showing us that our socio-political activities are governed not by the logic of reason, but, rather, by the logic of contrariety. In other words, while Habermas seems to envision the possibility of a final solution to socio-political conflict, Mouffe understands conflict as the very ontological given of socio-political reality. Any political change, therefore, for Mouffe, is conditioned not by agreement, but, rather, by disagreement. Nevertheless, despite their fundamental difference, I believe that both Habermas and Mouffe presuppose that language, in this case discourse, is part of the ontological makeup of socio-political interactions—hence, of democratic discourse. In consideration of these two contrasting positions, it seems to me that Lummis, while pegging radical democracy on common sense language, understands democratic discourse as not simply a methodical process of agreement (ala Habermas), but, rather, it is an aggressive articulation of disagreement with centralized power, as pointed out above. Similar to Mouffe, Lummis understands democracy as a subversive (antagonistic) stance, one that competes (agonistic) with a given oppressive order. Notwithstanding this fundamentally subversive character of democracy, Lummis notes that this subversive performance is a moral discourse since it is normatively informed by a common cause: the realization of freedom.

Since democracy is an ideal, a way of life, and not a method, it is not simply a specific kind of government or economic system (as opposed to the commonly held idea). On the contrary, democracy is the goal that a government or an economic setup should strive to achieve. G. W. F. Hegel is, of course, decidedly an anti-democratic philosopher, but, ironically enough despite his political elitism, the same idea he used in *Grundlinien der Philosophie des Rechts (Elements of the Philosophy of Right)*, *Sittlichkeit* or the “ethical life” life of a community, could also be used to describe what democracy is. This is precisely what we can observe Axel Honneth is doing in *Das Recht der Freiheit (Freedom’s Right)*, a point which I shall elaborate in the next section. Suffice it to say for now that democracy is a historical project that requires the active participation of a community of individuals. To quote Hegel:

---

<sup>20</sup> Cf. Jove Jim S. Aguas, Paolo A. Bolaños, and Jovito V. Cariño, “The Spectre of Terror: Philippine Democracy and the Threat of the New (Ab)normal,” in *Interfere Blog* (28 August 2020), <<https://interferejournal.org/2020/08/28/the-spectre-of-terror/>>.

<sup>21</sup> Laclau and Mouffe, *Hegemony & Socialist Strategy*, 126.

The *right of individuals to their subjective determination to freedom* is fulfilled in so far as they belong to ethical actuality; for their certainty of their own freedom has its *truth* in such objectivity, and it is in the ethical realm that they *actually* possess *their own* essence and their inner universality.

Those pedagogical experiments in removing people from the ordinary life of the present and bringing them up in the country (cf. Rousseau's *Emile*) have been futile, because one cannot successfully isolate people from the laws of the world. Even if young people have to be educated in solitude, no one should imagine that the breath of the spiritual world will not eventually find its way into this solitude and that the power of the world spirit is too weak for it to gain control of such remote regions. The individual attains his right only by becoming the citizen of a good state.<sup>22</sup>

According to Lummis, this participation, that is the integration of an individual into society must be measured against the backdrop of democracy:

... democracy is one of those beautiful, absolute, clear principles ... that poses a maddening, tantalizing puzzle to humankind. It is because there is no sure, fixed solution to this puzzle—the puzzle of how to realize democracy in our collective life—that our commitment to it can take the form only of a historical project. And how successful institutions may be in coming close to it, democracy itself—like justice, equality, and liberty—remains as a critical standard against which all institutions may be measured.<sup>23</sup>

As such, we can view democracy as a normative principle inasmuch as it is “a critical standard against which all institutions may be measured.” Lummis goes as far as saying that radical democracy “is the foundation of all political discourse”<sup>24</sup> inasmuch as it is concerned about the object of politics which is “power.” The normative content of democracy, which is freedom, is the

---

<sup>22</sup> G.W.F. Hegel, *Elements of the Philosophy of Right*, trans. by H.B. Nisbet (Cambridge: Cambridge University Press, 1991), §153.

<sup>23</sup> Lummis, *Radical Democracy.*, 22-23.

<sup>24</sup> *Ibid.*, 26.

principle behind the people's call for justice, that is to say, the eradication of injustice. In this sense, then, it appears to me that the agonal discourse that Mouffe has been describing is not incompatible with the notion of radical democracy that Lummis is proposing. However, what Mouffe does not clearly enunciate is the moral import of agonal discourse. One could, nevertheless, assume that, since antagonisms are political reactions to a dominant order (presumably with the tendency to oppress, like, martial rule), Mouffe is actually motivated by her critical outlook on oppressive socio-political orders.

### Critical Theory and the Idea of Freedom

I argue that the link between radical democracy and critical theory is the normative character of democratic discourse. Based on Max Horkheimer's *Critical Theory* essays, we could summarize critical theory as a philosophical and political position via three interrelated claims: (1) critical theory is a discourse that is normatively based on human affairs; (2) critical theory advocates the abolition of slavery and social injustice; and (3) critical theory decentralizes the discourse of emancipation from the proletariat to other social groups.<sup>25</sup> All three normative claims, I believe, intensify the idea of radical democracy so far described in the foregoing. This is possible if we recast these normative claims within the ambit of radical democracy.

The anthropological shift in critical theory, I argue, brings back democratic discourse to the people who constitute society. I have mentioned earlier that democratic discourse is informed by the ordinary language of the people; hence, normatively, it is a practice that takes shape based on the behavioral patterns of people in a given community.

The second normative claim of critical theory is the strongest, the abolition of slavery and social injustice, and one which resonates very well with the basic normative question of radical democracy: "What is justice?" I take this question to mean, what are we going to do with social injustice?

The third normative claim of critical theory, the decentralization of the emancipative impulse, can also be described as the democratization of emancipation. The early members of the Frankfurt School veered away from Georg Lukacs's over-valorization of the role of the proletariat class in his *History and Class Consciousness* and they viewed emancipation as a possibility for the greatest number of people or for other social groups. They saw the

---

<sup>25</sup> Cf. Max Horkheimer, *Critical Theory: Selected Essays*, trans. by Matthew J. O'Connell, et al. (New York: Continuum, 2002). For a more detailed discussion, see my "The Ethics of Recognition and the Normativity of Social Relations: Some Notes on Axel Honneth's Materialist Philosophical Anthropology," in *Suri: The Official Journal of the Philosophical Association of the Philippines*, 1:1 (2012), 15-24.



over valorization of the proletariat class as another form of social elitism that has the potential to exclude other social groups.

In his debate with Nancy Fraser, Honneth defends his position against the former's accusation that the latter's recognition theory is nothing more than moral psychology and culturalism. This means, for Fraser, that Honneth's recognition theory is not able to practically address sociological and economic issues. Honneth, however, disputes this and argues that recognition theory offers a normative basis for freedom and justice.<sup>26</sup> Seen in this light, it is perhaps possible to interpret Honneth's theory of theories of recognition and social freedom within the context of radical democracy, that is, a radical democracy informed by the three normative claims of critical theory I pointed out above. As a third option, that is between Habermas's deliberative democracy and Mouffe's agonistic democracy, a theory of social justice may be positioned at the middle inasmuch as it takes deliberative participation as an essential element of democratic discourse, on the one hand, and also seriously takes conflict or disagreement as a normative element of the discourse.

As with Lummis's position that radical democracy is a historical project, Honneth views democratic discourse as a distillation of recognitive struggles. For Honneth, recognitive struggles are the social bases for the possibility of intact recognitive relations. This could be understood in two levels. First, recognitive struggles are the normative bases for the development of a sense of ethical subjectivity within each individual. As such, these struggles develop within individuals a sense of social dignity inasmuch as they become part of the dynamics of the community. Second, recognitive struggles could also be manifestations of the clamor of disrespected individuals or social groups to regain their sense of social integrity. Following Hegel and Mead, Honneth notes, that "individuals must know that they are recognized for their particular abilities and traits in order to be capable of self-realization, they need a form of social esteem that they can only acquire on the basis of collectively shared goals."<sup>27</sup> As such, a claim can be made that a democratic recognitive process is tenable inasmuch as it aims to regain the social dignity of individuals or social groups. Said another way, the radicality of Honneth's recognition theory can be observed in its attempt to bridge the gap between, on the one hand, the experience of individual autonomy and bodily integrity of individuals, and, on the other hand, how these are made possible through the dynamics of community life.

---

<sup>26</sup> See Axel Honneth, "Redistribution as Recognition: A Response to Nancy Fraser," trans. by Joel Golb and James Ingram, in *Redistribution or Recognition? A Political-Philosophical Exchange* (London: Verso, 2003), 110-197.

<sup>27</sup> Axel Honneth, *The Struggle for Recognition: The Moral Grammar of Social Conflicts*, trans. by Joel Anderson (Cambridge, Massachusetts: The MIT Press, 1995), 178.

Meanwhile, in *Freedom's Right*, Honneth pursues Hegel's immanent approach of "picking up on values and ideas already institutionalized in society."<sup>28</sup> These institutionalized values, according to Honneth, must have a "moral" motivation that leads to the realization of "justice." Honneth contextualizes justice further by arguing that justice depends on the meaning ascribed to the idea of freedom. Nevertheless, he tries to veer away from a presentation of justice that is "merely formal" or "abstract."<sup>29</sup> Instead, Honneth attempts to provide a theory of justice that is already normatively connected to social reality. By social reality he means the "values and ideas already institutionalized in society." Honneth adds: "the demands of justice turn out to be the essence of the norms that contribute to the most appropriate and comprehensive realization of prevailing values within various different systems of action."<sup>30</sup> In this context, Honneth seems to be claiming that the content of norms (values and ideas instituted in society) is precisely what social justice demands—in other words, as with Lummis, the moral motivation of these norms is the realization of democratic life which is, I argue, nothing else but the realization of social justice. Against the backdrop of Hegel's critical stance against democracy's tendency towards an indefensible liberal individualism in the *Elements of the Philosophy of Right*, Honneth provides an alternative notion of individual freedom in *Freedom's Right*—offering not only an extension of Hegel's project but also a radical alternative to the latter's view on democracy.

In order to make the connection between democracy and social justice—and, therefore, the connection between democracy and critical theory—it is important that we have a working understanding of Honneth's idea of justice. In conceptualizing a concretely normative, as opposed to formal, notion of justice, Honneth argues that an idea of justice is only meaningful if understood in the context of freedom because, as an ethical idea, justice entails "the meaning that individual freedom takes on the differentiated spheres of action in accordance with their respective function."<sup>31</sup> This means that justice is demanded every time freedom is invoked and there is a plurality of instances of invocations. In *Freedom's Right*, Honneth identifies three conceptions of freedom which, for him, are various assumptions of what individual freedom entails: 1) negative freedom, 2) reflexive freedom, and 3) social freedom.

---

<sup>28</sup> Axel Honneth, *Freedom's Right: The Social Foundations of Democratic Life*, trans. by Joseph Ganahl (Cambridge: Polity Press, 2014), 63.

<sup>29</sup> *Ibid.*

<sup>30</sup> *Ibid.*, 64.

<sup>31</sup> *Ibid.*

The first, negative conception of freedom assumes that a legally protected sphere in which subjects can act on their own unreflected preferences is a crucial part of individual freedom ...

... reflexive idea claims that freedom depends on the performance of intellectual acts, which are nevertheless regarded as normal acts performed by every competent subject ...

... social idea of freedom takes account of additional social conditions, linking the realization of freedom to the condition that other, accommodating subjects confirm my own aims.<sup>32</sup>

Honneth's three conceptions of freedom is a recasting of the three spheres of "right" presented by Hegel in the *Elements of the Philosophy of Right*: (1) abstract right (*Recht*), (2) morality (*Moralität*), and (3) ethical life (*Sittlichkeit*). Abstract right refers to the "inherently individual will of a subject" or "exclusive individuality,"<sup>33</sup> pertaining to non-interference as a form of respecting individuals. Honneth recast this as negative freedom, that is, the freedom of subjects to pursue their unreflected interest without external interference. The moral sphere, according to Hegel, is the sphere of self-reflection and human interaction, wherein the subjective will becomes conscious of itself through the consciousness of the existence of other subjective wills. The "moral point of view," Hegel notes, is "the point of view of *relationship, obligation, or requirement*."<sup>34</sup> For Hegel, the establishment of "subjectivity" requires that the abstract will becomes concrete through an interaction with the objective world.<sup>35</sup> Meanwhile, for Honneth, the moral sphere is the domain of reflexive freedom, wherein the autonomy of an individual is demonstrated through the performance of "intellectual acts" that requires one's reflective comportment. Both the abstract and moral spheres are eventually superseded by what Hegel calls the sphere of ethical life, wherein subjective will is fully integrated in a community which is objective. Hegel maintains that, "Ethical life is ... the *concept of freedom which has become the existing ... world and the nature of self-consciousness*."<sup>36</sup> Moreover, the objective community transcends not only the abstract will but also the reflective will of morality, as the reflective will is still will unto others. Through the notion of ethical life, Hegel attempts to ground subjective

---

<sup>32</sup> *Ibid.*, 65.

<sup>33</sup> Hegel, *Elements of the Philosophy of Right*, §34.

<sup>34</sup> *Ibid.*, §108.

<sup>35</sup> *Ibid.*, §106.

<sup>36</sup> *Ibid.*, §142.

feelings of freedom in the normative objectivity of “*laws and institutions which have being in and for themselves.*”<sup>37</sup> Hegel expresses this idea alternatively as the “concrete concept of freedom,” wherein “we are not one-sidedly within ourselves, but willingly limit ourselves with reference to an other.”<sup>38</sup> Honneth provides an inflection of this Hegelian notion of concrete freedom as social freedom, that is, freedom conditioned by one’s membership in a community, wherein the experience of freedom presupposes that other free individuals recognize me as a member of the community, thereby recognizing also my peculiarities as an individual will. As a participating member of the community, an individual discovers mutual recognitive relations in various societal institutions, such as, family, economy, and politics. These institutions ensure that concrete freedom is realized, thereby achieving what has been hitherto referred to as justice.

### Freedom and Democracy

An ethical society, for Honneth, performs normative practices that maintain social reproduction. One such practice is “democratic politics.” I choose only here to discuss democratic politics as an aspect of the ethical society, first, because of space constraints, but, secondly and more importantly, because it is precisely an aspect of Honneth’s work that departs from a basic assumption of Hegel’s *Elements of the Philosophy of Right*, namely, that the ideals of an ethical society are opposed to democracy.<sup>39</sup> According to Honneth:

We will therefore have to depart from the model presented in Hegel’s *Philosophy of Right* in order to begin the normative construction of this third sphere, which in turn can only be analysed adequately if we understand it as an embodiment of social freedom: the institution of the democratic public or ‘public sphere’, a social space in which citizens form generally acceptable beliefs through deliberative discussion, beliefs that form the principles to be obeyed by the legislature in accordance with the rule of law.<sup>40</sup>

---

<sup>37</sup> *Ibid.*, §144.

<sup>38</sup> *Ibid.*, §7.

<sup>39</sup> For a more detailed discussion of Hegel’s critique of democracy, see W. G. Stratton, “The Problem of Democracy in Hegel’s Philosophy of Law,” in *Archiv für Rechts- und Sozialphilosophie*, 74:1 (1988), 33-41. For supplemental discussion, see Karin de Boer, “Democracy Out of Joint? The Financial Crisis in Light of Hegel’s *Philosophy of Right*,” in *Bulletin of the Hegel Society of Great Britain*, 66 (2012), 36-53.

<sup>40</sup> Honneth, *Freedom’s Right*, 254.

This is what I meant when I mentioned earlier that Honneth offers not only an extension of Hegel's political project but also a radical alternative to the latter's critique of democracy. Hegel's disapproval of democracy goes as far as leaving no room for public referenda or plebiscites in his ideal conception of the state—in other words, no chance for the less financially fortunate “rabble”<sup>41</sup> to participate in the affairs of the state. Hegel, moreover, is largely opposed to open franchise voting and, instead, proposes that the right to suffrage is restricted only to the members of the nobility and the middle class.<sup>42</sup> Meanwhile, Honneth distances himself from this Hegelian position and reinstates to the members of society the space for democratic deliberation, the “public sphere.” It is important to point out that Honneth does not view democratic deliberation in a purely juridical or proceduralist sense since deliberations of this sort have the tendency to “ignore or downplay the dependence of deliberative decision-making on ‘free’ conditions in the other constitutive spheres of society,”<sup>43</sup> namely, family (personal relationships) and economy (the market). It is in this context that we must understand Honneth's caveat that his theory of justice is normatively grounded in social reality, as opposed to abstract conceptions of justice. This charge of abstractness, I believe, is also aimed at Hegel's tendency in the *Elements of the Philosophy of Right* to propose a top-down theory of state organization, instead of a bottom-up or grassroots approach. As such, it would not be surprising if Hegel is viewed as a political elitist, a view that, unfortunately, dampens the more nuanced and critical aspects of his theory of freedom. By distancing himself from Hegel, at least on the issue of democracy, Honneth re-envisioning a more concrete idea of social freedom that is compatible with the fundamental principle of democratic life—he argues that, “the realization of social freedom in the democratic public sphere depends at the very least on the partial realization of the principles of social freedom in the spheres of personal relationships and the market.”<sup>44</sup> This is key in understanding Honneth's theory of democracy: democracy or “deliberative will-formation” is legitimate only when “it learns, in the process of continuous debate over the conditions of social inclusion, the necessity of supporting struggles for social freedom in the two other spheres.”<sup>45</sup> In other words, the social freedom that the public sphere promises is only derivative of the two bottom spheres of family and economics. This brings us back to what I mentioned earlier regarding common sense language as constitutive of democratic praxis—as the moral content of such praxis. Given the above

---

<sup>41</sup> Hegel, *Elements of the Philosophy of Right*, §244.

<sup>42</sup> *Ibid.*, §§305-311.

<sup>43</sup> Cf. Honneth, *Freedom's Right*, 254.

<sup>44</sup> *Ibid.*

<sup>45</sup> *Ibid.*, 255.

forage for the meaning of freedom and justice, I can now claim, through Honneth, that common sense language is normatively grounded in the dialectics of our individual right to be free and our intersubjective interactions. I also think that, by emphasizing the normative role of the dialectics of personal relationships and the market—of family and economy—Honneth is revivifying one of the most important insights of Marx regarding the social reproduction of class consciousness, where a group of individuals collectively develop an awareness of their concrete socio-political condition.<sup>46</sup> By going through this process of self-awareness and self-realization a group of individuals dialectically develop a sense of individual and group identity (class). Marx usually refers to this process as a “struggle” that is characteristic of a “deeper-lying antagonism” between two social groups, that is, between an oppressor and the oppressed.<sup>47</sup> I argue that the dialectical interaction between oppressor and oppressed marks a democratic moment in Marx’s theory of class consciousness, as it reflects the radical essence of democratic praxis, namely, a social group’s articulation of an antithesis to an oppressive group or system.

### **Concluding Note: The *Dêmos* and the Experience of Suffering**

In his 1940 essay, “On the Concept of History,” Walter Benjamin writes:

The tradition of the oppressed teaches us that the “state of emergency” in which we live is not the exception but the rule. We must attain to a conception of history that accords with this insight. Then we will clearly see that it is our task to bring about a real state of emergency, and this will improve our position in the struggle against fascism. One reason fascism has a chance is that, in the name of progress, its opponents treat it as a historical norm. The current amazement that the things we are experiencing are “still” possible in the twentieth century is *not* philosophical. This amazement is not the beginning of knowledge—unless it is the knowledge

---

<sup>46</sup> Karl Marx discusses his theory of class consciousness in *Capital: A Critique of Political Economy*, Volume I, trans. by Samuel Moore and Edward Aveling, in *Marx & Engels Collected Works* (Great Britain: Lawrence & Wishart, 2010). Georg Lukács offers perhaps the most comprehensive commentary on Marx’s theory of class consciousness in *History and Class Consciousness: Studies in Marxist Dialectics*, trans. by Rodney Livingstone (Cambridge, Massachusetts: The MIT Press, 1971).

<sup>47</sup> Marx, *Capital*, 146.



that the view of history which gives rise to it is untenable.<sup>48</sup>

Benjamin apparently paints a foreboding image of history. History is not a smooth and linear progression towards the future, but, rather, is marred by the tumultuous dialectics of the past and present. There is a tendency, Benjamin laments, for us moderns to forget about the sufferings of the past in the name of progress; by forgetting the past, the histories and identities of those who suffered are lost. This is the tragedy of the present which Benjamin calls the “state of emergency,” a state which we are not yet fully aware of—hence, we have lost so much of our histories and of ourselves. Because we continue to forget, we are not only complicit to the crime of fascism, but also continue to suffer. As opposed to a progressive historical materialism, what Benjamin offers as an alternative is a historical materialism from the standpoint of suffering or, in his words, “the tradition of the oppressed.” As an image of thought, the tradition of the oppressed provides a historical lens that serves as a constant reminder of past sufferings and warning about the possibility of future sufferings. In this context, we could view this image of thought as a materialist—that is to say, concrete—ethical starting point. I believe that this is what Benjamin means when he says, “We must attain to a conception of history that accords with this insight,” so that we “will improve our position in the struggle against fascism.” In this sense, if we understand Marx’s insight on the self-realization of consciousness, a historical point of view informed by the reality of suffering becomes the ground for the self-realization of a consciousness that is still suffering. It is suffering because it is still unjustly dominated. The tradition of the oppressed widens our cognitive field and, hopefully, also our vocabulary so that we are going to be better equipped in our struggle against any form of fascism. It is also in this context that we are able to understand what Honneth reminds us about social freedom, that it is not achievable in the consciousness of those who oppress, but, rather, in the consciousness of those who are oppressed.

In my attempt to bring together normative elements from critical theory and the promise of radical democracy, I used, as a point of departure, a working definition of radical democracy offered by Lummis. I am convinced that Lummis offers a powerful description of the radicality of democratic praxis, inasmuch as he emphasizes that the essence of such praxis is precisely the articulation of an antithesis to centralized power. I tried to elaborate on the praxis or articulation of democratic discourse by pointing out three models: (1) Habermas’s deliberative discourse, (2) Mouffe’s

---

<sup>48</sup> Walter Benjamin, “On the Concept of History,” in *Walter Benjamin: Selected Writings, Volume 4, 1938-1940* (Cambridge, Massachusetts: Harvard University Press, 2003), 392.

agonistic discourse, and (3) Honneth's dialectics of social freedom. While I find in Honneth, who represents the contemporary strain of critical theory, a possible rapprochement between the deliberative and agonistic discourses, I think that Honneth's model is the one that is most committed to the normative claims of critical theory, especially the claim on the abolition of slavery and social injustice. Inspired by Hegel's idea of *Sittlichkeit*, Honneth advances a theory of justice that is grounded in a concrete notion of social freedom which is, at the same time, decidedly anti-Hegelian since Honneth understands the dynamics of social freedom as essentially democratic. As a normative practice, democratic politics provides an agonal space both for -self-realization and societal inclusion. The radicality of Honneth's theory of democracy is found in his inflection of Marx's description of the dialectical self-realization of consciousness in the face of conflict. We can relate the Benjaminian image of the tradition of the oppressed with Honneth's commitment to a critical theory grounded in the normativity of the experience of social injustice or suffering. Whose suffering is it that becomes the normative ground for democratic praxis? I believe it is the suffering of the *dêmos*.

*Department of Philosophy, University of Santo Tomas, Philippines*

## References

- Aguas, Jove Jim S., Paolo A. Bolaños, and Jovito V. Cariño, "The Spectre of Terror: Philippine Democracy and the Threat of the New (Ab)normal," in *Interfere Blog* (28 August 2020), <<https://interferejournal.org/2020/08/28/the-spectre-of-terror/>>.
- Aristotle, *The Constitution of Athens*, trans. by Thomas J. Dymes (London: Seeley and Co., Limited, 1891).
- Benjamin, Walter, "On the Concept of History," in *Walter Benjamin: Selected Writings, Volume 4, 1938-1940* (Cambridge, Massachusetts: Harvard University Press, 2003).
- de Boer, Karin, "Democracy Out of Joint? The Financial Crisis in Light of Hegel's *Philosophy of Right*," in *Bulletin of the Hegel Society of Great Britain*, 66 (2012).
- Bolaños, Paolo A., "The Ethics of Recognition and the Normativity of Social Relations: Some Notes on Axel Honneth's Materialist Philosophical Anthropology," in *Suri: The Official Journal of the Philosophical Association of the Philippines*, 1:1 (2012).
- Cahiles, Gerg, "Politics of recognition: The power of religious endorsement in the 2019 polls," in *CNN Philippines* (6 April 2019), <<https://cnnphilippines.com/news/2019/4/10/religious-endorsement-philippine-elections.html>>.

- Habermas, Jürgen, *Moral Consciousness and Communicative Action*, trans. by Christian Lenhardt and Shierry Weber Nicholsen (Cambridge: Polity Press, 1990).
- \_\_\_\_\_, *The Theory of Communicative Action*, Vol. 1, trans. by Thomas McCarthy (Boston: Beacon Press, 1984).
- Hegel, G.W.F., *Elements of the Philosophy of Right*, trans. by H.B. Nisbet (Cambridge: Cambridge University Press, 1991).
- Honneth, Axel, *Freedom's Right: The Social Foundations of Democratic Life*, trans. by Joseph Ganahl (Cambridge: Polity Press, 2014).
- \_\_\_\_\_, "Redistribution as Recognition: A Response to Nancy Fraser," trans. by Joel Golb and James Ingram, in *Redistribution or Recognition? A Political-Philosophical Exchange* (London: Verso, 2003).
- \_\_\_\_\_, *The Struggle for Recognition: The Moral Grammar of Social Conflicts*, trans. by Joel Anderson (Cambridge, Massachusetts: The MIT Press, 1995).
- Horkheimer, Max, *Critical Theory: Selected Essays*, trans. by Matthew J. O'Connell, et al. (New York: Continuum, 2002).
- Laclau, Ernesto and Chantal Mouffe, *Hegemony & Socialist Strategy: Towards a Radical Democratic Politics* (London: Verso, 1985).
- Lukács, Georg, *History and Class Consciousness: Studies in Marxist Dialectics*, trans. by Rodney Livingstone (Cambridge, Massachusetts: The MIT Press, 1971).
- Lummis, C. Douglas, *Radical Democracy* (Ithaca: Cornell University Press, 1996).
- Maboloc, Christian Ryan, "President Rodrigo Duterte and the Birth of Radical Democracy in the Philippines," in *International Journal of Politics and Security*, 2:3 (May 2020).
- Marx, Karl, *Capital: A Critique of Political Economy, Volume I*, trans. by Samuel Moore and Edward Aveling, in *Marx & Engels Collected Works* (Great Britain: Lawrence & Wishart, 2010).
- Mouffe, Chantal, "Deliberative Democracy or Agonistic Pluralism?," in *Social Research*, 66:3 (Fall 1999).
- \_\_\_\_\_, *The Return of the Political* (London: Verso, 1993).
- Nicolas, Fiona, "Duterte, Marcos get INC endorsement for May 9 polls," in *CNN Philippines* (5 May 2016), <<https://cnnphilippines.com/news/2016/05/05/Iglesia-ni-Cristo-endorsement-duterte-marcos-May-9-elections.html>>.
- Stratton, W. G., "The Problem of Democracy in Hegel's Philosophy of Law," in *Archiv für Rechts- und Sozialphilosophie*, 74:1 (1988).

Article

## Moral Cognitivism: 'Motivation' and Agency

---

**Bob Brecher**

**Abstract:** Moral cognitivism is pointless unless what people know they ought to do leads them to do it. But then, how can we act unless we're motivated to act? So, moral knowledge requires motivation if it is to be realized, and because we need to want to act morally in order to do so, cognitivism has disappeared. I shall attempt to deal with this problem by proposing that we jettison the very idea of motivation. Taking my cue from Richard Norman's "Practical reasons and the redundancy of motives,"<sup>1</sup> and the direction of some of the work that precedes it, I shall suggest that rational action has no need of any notion of motivation at all. Reasons, I shall argue, are all that we need to explain not only belief, but also action, since agency is not something that requires to be "switched on"; rather it is integral to personhood. "Motivation," "will" and all that goes with them can be simply dropped as no more necessary in the context of action than that of belief.

**Keywords:** action, morality, motivation, reasons

### 1. Introduction

**W**hy is motivation such a problem for moral cognitivists? Because moral cognitivism is pointless unless what people know they ought to do leads them to do it, to act on that knowledge; and to act appears to require that we are motivated to act. There is nothing necessarily odd about knowing that Quito is the capital of Ecuador but doing nothing in light of that knowledge. But if you knew it was wrong to torture a prisoner, it would be bizarre if you did it nonetheless. As Jean Hampton puts it in her still definitive account of reasons and reason, the view that there is no "logical connection between an action's moral characteristics and a reason

---

<sup>1</sup> See Richard Norman, "Practical Reasons and the Redundancy of Motives," *Ethical Theory and Moral Practice* 4 (2002), 3–22.

for performing it”<sup>2</sup> empties morality of what is fundamental to it, namely, either that something needs to be done, or that it must not be done. Without a plausible link between moral knowledge and moral action— that is to say, an account of how moral beliefs and judgments lead us to act—any analysis of morality is empty. On this, Hume was entirely right. Again, as Hampton asks: suppose Albert knows it is wrong to burn cats but that this knowledge does not impinge on what he does, such a view

not only separates Albert’s knowledge that cat-burning is wrong from a motive not to engage in it; it also fails to locate any way in which Albert has a reason not to burn cats. Albert’s labelling the activity “wrong” is not only without intrinsic motivational impact ... it is also without any authoritative impact. Albert is free to say, ... “Okay, cat burning is wrong, but what does that label have to do with whether or not I should do it? I want to do it, so I have no reason not to do so.” In other words, there is no way to say that Albert has a reason not to burn cats if, as it happens, he would like to do so.<sup>3</sup>

So, it is difficult to see what Albert actually knows: in asking why he should take any notice of the moral reasons he knows there are not to burn cats, he can’t mean exactly what he appears to be asking. Nor does he. What he is actually asking is why he should take any notice of what other people take to be moral reasons; and that is an entirely different question. For there being a reason is not the same as having a reason, as I shall go on to discuss. To acknowledge a reason as a moral reason is precisely to acknowledge it as a reason to do or not to do something. This isn’t peculiar to moral reasons: it applies to practical reasons quite generally. To acknowledge the pouring rain as a reason to take an umbrella is to acknowledge that one ought to do so— other things being equal, since a reason in this context need not, of course, be a decisive reason. That is how everyday practical reasons differ from moral reasons; but that does nothing to undermine the practicality of either of them. Their “action-guidingness,” as one might say, remains in place. Nor does this differentiate theoretical reasons from practical reasons. Knowing that India is larger than Malaysia, and that China is bigger than India, is a reason for knowing that China is bigger than Malaysia. If I were then to ask why I should believe this, I would simply have failed to grasp what a reason was. The

---

<sup>2</sup> Jean Hampton, *The Authority of Reason* (Cambridge: Cambridge University Press, 1998), 101. This is, of course, Thrasymachus’ notorious question in Plato’s *Republic* about why he should not do whatever he wants if he can get away with it.

<sup>3</sup> *Ibid.*

“belief-guidingness” of theoretical reasons, whether logical or empirical, is exactly parallel to the “action-guidingness” of practical reasons. Indeed, one could regard believing something as one sort of action: certainly, acquiring beliefs is not something passive, not something like, say, becoming aware of the sound of a car going down the road.

But this raises an immediate difficulty. We all seem sometimes to know perfectly well what we ought, morally, to do, and yet fail to do it nevertheless. Why? The standard account—the belief-desire model of motivation, derived from Hume—offers an obvious answer: we don’t want to (enough). We have to want to do something if we are to do it, whether in a moral context or otherwise. My wanting to go to the bar is a reason for my going just because I want to: the desire is built into the reason. And in the moral case, so the story goes, where what I want, rather than what I ought, to do, appears not to come into it, I in fact do what I ought only if I want to do so at some “higher” level; if I want to be a good sort of person more than I want not to bother helping so-and-so. The story goes that the required motivation is sparked off, or, at any rate, somehow initiated, by desire, by what I want at some level or another. We have to want, if not directly to do something, then at least the consequences of doing it; we might not want to visit an aged relative in the hospital, but we do want her to think well of us, and so we go.

This sort of account—a determinedly psychologized account of morality—has it that, since both reason and the material world are motivationally inert, motivation—what moves us—has to come, so to speak, from within us. And only our wants, more or less broadly conceived, fit the bill. Given that we want something, and in light of the information we have about how to get it—our beliefs—it is our wanting it that gets us going, that motivates us. We have at some level to want to do what we do. But this is fundamentally mistaken. Moral reasons hold good quite regardless of what we want.<sup>4</sup>

That is what Kant meant by insisting that morality’s claims are categorical, and not hypothetical: morality requires that we do what is right. He insists that we ought to do what is right just because it is right. So, we have to be able to act, as Kant puts it, from reasons alone. To quote Hampton again:

Kant’s defense of the moral law is of this form: The moral law provides reasons for us, and is ultimately authoritative for us, but its authority does not in any way

---

<sup>4</sup> See Bob Brecher, *Getting What You Want?: A Critique of Liberal Morality* (London & New York: Routledge, 1997).



*depend* on its motivational efficacy, although it is true that, by virtue of its authority, the moral law is (necessarily) not only motivationally efficacious but sufficient to move us.<sup>5</sup>

So, the problem appears to be this: how to offer an account of how moral reasons can be “motivationally efficacious,” how they can be “sufficient to move us,” let alone how, on Kant’s view, they cannot fail “to move us.” But this appearance is misleading: the questions raised are themselves mistaken. That they lead inevitably to a rejection of moral cognitivism is not why they are mistaken, of course; that would be an obvious mistake. So why are they mistaken?

## 2. “Motivation,” reasons and actions: some difficulties

In this explicitly programmatic discussion, I’m going to propose that we need to jettison not only wants, but also the very idea of motivation. Taking my cue from Richard Norman’s oddly overlooked 2002 paper, “Practical reasons and the redundancy of motives,”<sup>6</sup> and the direction of some of the work that precedes it, I want to suggest that it is not that reasons motivate without help from any sort of affect, but rather that rational *action has no need of any notion of motivation at all*. We can say all we need to say about why and how we do things without any talk of motive or motivation at all. In short, it is a mistake to think that some psychological state of affairs has to be obtained—my being motivated—if the reasons I have for doing something are to lead me—thanks to my being in such a state—to do it. I take this view because it is the only way I can see of avoiding the conclusion that we have to want to do what is right if we are to do it: we need to get rid of the very notion of motivation and to show that there being good reasons to do something is enough to account for our doing it. We need nothing *further* by way of positing some affective state of affairs to do what there is (sufficient) reason to do: that is to say, that “we do not need a theory of motivation at all,” as Norman puts it.<sup>7</sup> Or to put it another way: the holiday that European language has been on since Augustine invented “the Will” as a means of dealing with theodicy is one that it should never have taken.

---

<sup>5</sup> Hampton, *op. cit.*, 68, n. 29. I think that Hampton is misled here by the idea of motivation when she goes on to say, in the same footnote, that Kant “would agree that moral reasons, as it happens, are also motivational”: for as she herself has just made clear, there is no contingency about that. See also fn. 24.

<sup>6</sup> Richard Norman, “Practical Reasons and the Redundancy of Motives.”

<sup>7</sup> *Ibid.*, 3.

What exactly is it that is said to require that a person be motivated if they are to be able to act? A theory of motivation “seems to be intended to fill a supposed gap between ‘reasons’ and ‘action.’ How, it may be asked, can our having *reasons* to act result in *action*?”<sup>8</sup> Interestingly, we don’t ask how our having reasons to believe that such-and-such is so results in belief. But why do we take this for granted? Why is it taken to be just obvious that we need to be motivated to go for a walk, but do not need to be motivated to believe that it is raining? Well, presumably because actions are one thing and beliefs and judgements quite another: very roughly, actions are physical; beliefs and judgements, mental. And reasons, being mental entities, thus have some sort of “direct connection” with beliefs that they don’t have with actions. For actions consist in physical movements and are thus objects of scientific explanation; they require to be capable of being explained in causal terms. Actions need to be “prompted” by something which is also material. Reasons, however, since they are not material but mental entities, cannot fit the bill: “[B]elieving that I have a reason’ is one thing, it may be said, ‘acting’ is another, and it is the supposed gap between the two that a theory of motivation is supposed to fill.”<sup>9</sup> But then, of course, once we have a scientific account, an explanation, we have neither need of, nor space for, any other. Once explained, the action is already intelligible so that reasons fall out of the picture altogether—unless, as Davidson famously argues, they are in fact a variety of cause in the first place.<sup>10</sup> But then it turns out that they are not really *reasons* at all, inasmuch as something has already to figure in our motivational psychology if it is to be a reason. Reasons have to be “our” reasons; they are reasons “for us,” something we “have.” In short, the account given of reasons on this neo-Humean picture is a thoroughly psychologized one: as Dancy aptly puts it in one of his earlier treatments, “[The] general form of this position is that with Humean reasons in place, we need no other sort of reasons than Humean ones. Or, more strongly, there is *no space left* for any other sort of reason.”<sup>11</sup> Once we know that such-and-such is a reason “for you” to go to the bar, we have all the explanation we need, indeed all the explanation there could be, of why you went to the bar. “There is no space left” to ask whether or not there was, in fact, a reason for you to go; what you did is already fully intelligible in light of relevant facts *about you* (set in the appropriate context).

---

<sup>8</sup> *Ibid.*

<sup>9</sup> *Ibid.*

<sup>10</sup> Donald Davidson, “Actions, Reasons, and Causes” reprinted in his *Essays on Actions and Events* (Oxford: Oxford University Press, 1980), 3–20.

<sup>11</sup> Jonathan Dancy, “Why There Is Really No Such Thing as the Theory of Motivation,” in *Proceedings of the Aristotelian Society*, 95 (1995), 10.

What is perhaps most significant about how reasons are understood on this view—a view which most of the Anglo-American tradition in philosophy has come largely to take for granted—is that they are essentially private, rather like pain. If such-and-such is a reason “for you,” then that is all there is to it; your sincere claim that it is “mine” cannot be challenged, however odd it might be. Reasons figure primarily as explanations and only contingently as justifications. That is why Davidson’s view of reasons as, basically, causes is so widely accepted: it is the tradition’s natural view. Of course, your believing that such-and-such is a reason to go to the bar explains your going whether or not it actually is a reason, whether or not it justifies your going. But believing that there is a reason and there being a reason are different states of affairs. Unhappily, the way in which we speak possessively of reasons obscures this. Reasons are—in some sense, and however difficult to explicate—a feature of the world, not of the psychological constitution of any individual. They are not “yours,” “mine” or “ours.” But the antipathy towards reasons being understood in this way, as “external,” runs deep. So, for example, Bernard Williams influentially argues that

Should we suppose that, if genuine external reasons were to be had, morality might get some leverage on ... the fanatical Nazi? ... I cannot see what leverage it would secure: *what would these external reasons do to these people, or for our relations to them?*<sup>12</sup>

But this misses the point entirely. The point of moral reasons is not whether, as a matter of contingent psychological fact, they would actually convince particular people: they might or might not. Fanatical racists are of course unlikely to listen; or, listening, be unable to grasp what is being said; or, able to grasp it, mistakenly suppose otherwise. Nor are they alone, of course. Furthermore, and turning to Williams’ own contrast with moral thinking, scientific thought, the situation is no different. The fact that many eminent scientists were unable or unwilling to listen to what Galileo was telling them, or, listening, remained unconvinced, makes not the slightest difference to the fact that there were reasons to suppose that the Earth went around the Sun. It makes no difference at all that, until Galileo made his

---

<sup>12</sup> Bernard Williams, “Replies,” in *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*, ed. by J.E.J. Altham and Ross Harrison, (Cambridge: Cambridge University Press, 1995), 216. Compare how he fails to see the same point on p. 215, perhaps because his opponents too write of reasons as something people “have”: “What is gained, except perhaps rhetorically, by claiming that A has a reason to do a certain thing, when all one has left to say is that this is what a *phronimos*, a decent person, or some such would do?” What a decent person would do—assuming that such a person would be one who did what was morally right—is what there was (sufficient) reason for them to do.

observations, no one “had” reasons to think so. There always had been such reasons (at least since the inception of the Solar System).

The baleful influence of this psychologized notion of reasons is widespread. Even thinkers who would not at all agree with Williams’ neo-Humean conception of what reasons are often fail to be sufficiently rigorous in their non-psychologism. For example, Geoffrey Sayre-McCord rightly points out, in the course of defending coherentism, that there is a great difference between someone’s being “justified in holding a particular belief” and the question of whether or not “the belief [is] justified”:<sup>13</sup> “one might justifiably believe what (as it happens) is false.”<sup>14</sup> I might, for instance, have good reason to believe that the temperature of the spiced sugar syrup I am making is 110 degrees because that is what it says on my hitherto entirely reliable sugar thermometer, even though it actually is not, because as it turns out, and unbeknownst to me, the thermometer is not working properly. And Sayre-McCord is right to go on to say that therefore, on his view, “a person’s belief is justified only if it coheres well with her other beliefs; whether it does is independent of whether she thinks it does.”<sup>15</sup> But something crucial is missing here: are her other beliefs *right*? It is not a question of whether or not she is justified in believing what she believes: as Sayre-McCord, says, she might be justified in believing something false. Rather, it is a question of whether or not she is *right* to believe what she believes. Otherwise, Williams’ and others’ “logically consistent racist” could be an exemplar of a coherentist view of morality. The question is not what reasons someone *has*—they could be false despite being justifiably held—but what reasons *there are*. The distinction is crucial.

### 3. Reasons, beliefs, and actions

Nor is the relation of reasons to action so very different from their relation to belief. Notice, first, that to make a judgement is to *do* something; it is not passive, not an event: it is not something that happens to us, that comes to us from outside, but something that we *make*. Our acquiring beliefs need not always be a matter of our imbibing and absorbing them from those around us, in the way that children might start to form their beliefs; or rather, might come to have their beliefs formed by people around them. Typically, as we grow up, so we exercise increasing discretion in what we believe: we learn to think for ourselves. And thinking is something that we *do*; it is not

---

<sup>13</sup> Geoffrey Sayre-McCord, “Coherentist epistemology and moral theory,” in *Moral Knowledge?*, ed. by Walter Sinnott-Armstrong and Mark Timmons, (Oxford: Oxford University Press, 1996), 146.

<sup>14</sup> *Ibid.*, 140.

<sup>15</sup> *Ibid.*, 177.

something that happens to us. Thinking is a species of activity. And in that case, why assume that acting requires that we be motivated—by desire or the will (whatever that may be)<sup>16</sup>—when desiring and willing *themselves* do not require motivation? Let me put it crudely, so as to make clear what is fundamental here. The neo-Humean response, or at least, a reconstruction of what I take the underlying thought to be, is that an “act of the mind” is not “really” an act. It is not something physical, and so only metaphorically an action. But if that is so, then, we have a puzzle. If to want, or to will, something is simply a mental state of affairs or, in the case of willing, a mental act, then, how do either of these serve to motivate (physical) action? Presumably, again, it must be that the similarity in respect of act outweighs the difference in respect of physical and mental. Such an account might well seem plausible: to *have* a reason is not an action, and there being a reason even less so. That is why the presence alone of a reason to act is insufficient, and requires the addition of affect, which is something—at least metaphorically, if not necessarily literally—active. Well, suppose that this is indeed plausible: but even if the “passivity” of reasons were enough to explain why they are unable to initiate action, the question would remain of exactly how reasons could then initiate thinking, deliberating, judging, and so on, since these are actions, even if “mental” actions. And if that is denied—if it is insisted that actions be understood in purely physical terms, as the movement of neurons or whatever—then, it is even more puzzling how reasons are able to initiate these, but not other sorts of “physical movement,” namely, actions. In short, the very idea of a reason becomes increasingly tenuous: on this sort of view, it becomes very hard to see what part reasons might have even in our deliberations, let alone our actions.<sup>17</sup>

To put it crudely: if “motivation” designates something “mental,” then how does it connect with actions? If, on the other hand, it designates something “physical,” then, how does it connect with reasons? As Hampton puts it, the difficulty on this view is how to explain that while the

authority (of reasons) is not the same as motivational efficacy, [but] it does seem as if there is some kind of link between the two. Even if one is not some kind of motivational internalist, one will likely still believe that

---

<sup>16</sup> A thorough-going defence of moral cognitivism would of course require not merely that we junk “motivation,” but also that we reject the very idea of “the will” and of “willing.” But that is for another occasion.

<sup>17</sup> Again, a defence of moral cognitivism will eventually require an account of the relation between something’s being the case and the reason, or reasons, why it is the case. The metaphors to which we have become accustomed—something “follows from” something else; reasons “impel” or “lead” us in various ways; it is the “force” of reasons that does the connecting work—are all of them less than perspicacious.

it is possible for human beings to be motivated by reasons.<sup>18</sup>

But how? For, as Hampton's Humean has it, "it doesn't make sense to say that it follows from the concept of having a reason to x that ... one also has a motive to x. Whatever our concepts are, they don't have control over the workings of the physical world."<sup>19</sup> Let us consider this in some detail.

First, it is clearly the case that my thinking about a physical object will not move it. However, if there is a reason for me to do something, and I do it, then, that makes a difference in the physical world. So, reasons, since they are a necessary component of action, clearly *impact* on the physical world, even if they do not control it. How action, as contrasted with behaviour, is possible is of course, a notorious difficulty. (I am not going to solve the free will/determinism issue here.) But the point is that *if* action is possible—*if* action is action and not merely behaviour—then, reasons figure *somehow* in the physical world. Again, Hampton's characterization is apt as an initial approach to the issue:

[W]hen we say that an agent acts "on," or "for the sake of" a reason, we are trying to say something about how this agent is "lured" to the action *by* the reason, as opposed to being driven or pushed into the action by some inner motivational force.<sup>20</sup>

Desires, passions, and other instances of affect, we might say, push us; reasons, by contrast, pull us. It is in that sense that we have control over our actions. To put it metaphorically, the pull of reasons can be accepted or resisted in a way that the push of what we want seems to admit of "no such choice or decision,"<sup>21</sup> inasmuch as our wants are part of the physical world. But though this way of putting it may initially seem helpful, the "push-pull" metaphor allows, in the end, too close a similarity between how affect works and how reasons work. The metaphor is too mechanical. For "[T]he practical force of normative reasons is not something that *makes* us act."<sup>22</sup> Rather, "... the practical force of reasons ... is *sui generis*, it is just the force of *reasons*, and cannot be equated with some other kind of force ..., something rooted in our

---

<sup>18</sup> Hampton, *The Authority of Reason*, 91.

<sup>19</sup> *Ibid.*, 70.

<sup>20</sup> *Ibid.*, 92.

<sup>21</sup> *Ibid.*

<sup>22</sup> Richard Norman, "Public Reasons and the 'Private Language' Argument," *Philosophical Investigations*, 23 (2000), 310.



social nature.”<sup>23</sup> Perhaps so. Certainly, what is centrally important is the distinction, in Hampton’s words, “between, on the one hand, reasons that are compelling considerations on which we choose to act, and on the other hand, nonrational motivators as mere ‘drivers’ of action....”<sup>24</sup> Except that there is considerable ambiguity about what exactly she means here. First, if the “nonrational motivators” which are “mere ‘drivers’ of action” are purely affect, then, there is no action at all, but only behaviour. Nor is there any action if “acting from desires” is understood as acting purely in response to, or on account of, desires; again, to “act” thus is to behave, and not to act at all. Only if my desires constitute the reasons for what I do, do I act “from desires.” Kant, for example, insists that the difference between a moral action and a non-moral action lies precisely in this: desires, or wants, cannot constitute (any part of) my reasons for acting if my action is to be a moral one:

I cannot have respect for inclination as such, whether it is mine or that of another; I can at most in the first case approve it and in the second sometimes even love it, that is, regard it as favorable to my own advantage. Only what is connected with my will merely as ground and never as effect, what does not serve my inclination but outweighs it or at least excludes it altogether from calculations in making a choice—hence the mere law for itself—can be an object of respect and so a command.<sup>25</sup>

Second, it does not follow at all “from the *concept* of having a reason” that “one also has a motive.” Rather, to say that I am motivated to do something is just a traditional but wholly misleading way of saying that there is a reason—one I recognize—to do it. The mistake is to suppose that there is something “behind” reasons, motivation, that is needed to theorise; and to suppose that, therefore, there really is an issue about how it can be, as Kant puts it, “that my psychology is under pressure by logic.”<sup>26</sup>

---

<sup>23</sup> *Ibid.*, 310.

<sup>24</sup> *Ibid.*, 293.

<sup>25</sup> Immanuel Kant, *Groundwork of the Metaphysics of Morals*, trans. by Mary Gregor (Cambridge: Cambridge University Press, 1997), 13 [4:400]. Cf. Kant, *The Metaphysics of Morals*, trans. by Mary Gregor (Cambridge: Cambridge University Press, 1996), 13 [6:213]: “That choice which can be determined by *pure reason* is called free choice. That which can be determined only by *inclination* (sensible impulse, *stimulus*) would be animal choice (*arbitrium brutum*). Human choice, however, is a choice that can indeed be *affected* but not *determined* by impulses, and is therefore of itself ... not pure but can still be determined to actions by pure will. *Freedom* of choice is this independence from being *determined* by sensible impulses; this is the negative concept of freedom. The positive concept of freedom is that of the ability of pure reason to be of itself practical.” In the end, even Kant is too much of a voluntarist.

<sup>26</sup> *Ibid.*, 69.

#### 4. Agents and agency

The empiricist tradition from Hobbes onwards has propagated an oddly mechanical notion of the individual, and thus, of human agency as an “extra” that requires explanation.<sup>27</sup> However, it is a picture of human agency on which no coherent understanding of morality is possible. For if teleological accounts are ruled out as accounting for actions because they are not scientific, then, *reasons* for action are ruled out too; and if there can really be no explanatory reasons for action, but only causal accounts of it, then, there are indeed no actions, only behaviour. If we understand ourselves as fundamentally physical entities that need a “shove” to get them moving, then no contortions, however ingenious, can make intelligible the notion of acting for reasons. No wonder, then, that on the neo-Humean picture, nothing finally remains of practical reason, let alone of morality—or even of agency.

The main reason why a need for some theory of motivation is assumed is that agency is thought of as something *exceptional*: the machine, normally static, needs something to get it moving. The assumption is that, as people, we *sometimes act*: “[T]he picture is of humans beings as essentially static, beings whose natural condition is one of inaction, and who need some kind of motivating force to effect the transition from inaction to action.”<sup>28</sup> But why not *start* with agency? Why not think of ourselves primarily as agents, who *sometimes do not act*? After all, we are as much “doing beings” as “thinking beings”; and as I have suggested, to think is to do something. We are not static objects which are, from time to time, “moved” to act. Rather, we are, first and foremost, agents: it is our doing things—whether physically, mentally, or both—that makes us persons and not just members of the species, *homo sapiens*. So, for instance, Martin Hollis points out that “it is not plainly true for everyone that all effort is at a cost. The idea that the rational agent always needs an incentive to *do anything at all* is dubitable.”<sup>29</sup> Or as Norman succinctly puts it, we need to remind “ourselves that, in the relevant sense, we are always acting, that is, for as long as we are conscious, and not asleep, under a general anaesthetic, comatose or dead.”<sup>30</sup>

---

<sup>27</sup> Thus, Hobbes said “For seeing life is but a motion of Limbs, the beginning whereof is in some principal part within ...” See Thomas Hobbes, *Leviathan* (Harmondsworth: Penguin, 1968), 81.

<sup>28</sup> Norman, “Practical Reasons and the Redundancy of Motives,” 7.

<sup>29</sup> Martin Hollis, *The Cunning of Reason* (Cambridge: Cambridge University Press, 1987), 19. He goes on to remark that this idea “is connected with the thought that voluntary social relations are instrumental and so engaged in only for mutual gain, which is dubitable too.” Altruism is impossible on the neo-Humean view of agency; it would always be an “altruism” at best, tinged by the desire to be altruistic, for on this view to “call an agent rational is to say merely that he reasons correctly in identifying the action likeliest to satisfy his preferences.” *Ibid.*, 60.

<sup>30</sup> Norman, “Practical Reasons and the Redundancy of Motives,” 8.

A great advantage of this perspective is that it enables us readily to distinguish two questions which are otherwise easily run together and which sow confusion in debates about moral cognitivism. One is the question misleadingly raised when in “the grip of this ‘picture which holds us captive’”:<sup>31</sup> “Why do you act (rather than not acting)?” But this is an unanswerable question, and thus, not properly a question at all.<sup>32</sup> To act is what agents just do, in the same way as thinking is something thinkers just do. The other question is the genuine, everyday question we happily ask when we ask why someone is doing *that*. And unlike the former, this is a perfectly good question. The point is that “Why are you doing that?” and “Why are you doing anything at all?” are profoundly different, notwithstanding any grammatical similarity.<sup>33</sup> Again, we need to *start* with what we are, namely rational agents. If we do that, then, the puzzle—about how something both static and non-physical, or reasons, can “move” us to do things—dissolves. We are freed from the assumption that some mediation is required between thought and action; and from the assumption that our “normal” state is a passive one. We no longer have to suppose that we have to invoke something special, something “out of the ordinary,” to explain why we are, unusually, moved to do something. Again, acting turns out to require no explanation at all, for, as agents, acting is just what we do:

If we are always acting, then it is misleading to talk about practical reasons as “reasons for acting.” Practical reasons are always reasons for doing *this*—for performing one specific action rather than another. *There are no reasons for acting, there are only reasons for actions.*<sup>34</sup>

We do not need something other than, or additional to, reasons for a particular action in order to act on those reasons: “people perform actions,” Norman says, “because they take themselves to have reasons for performing those actions, and ... their reasons explain why they act *as they do*.”<sup>35</sup> Of

---

<sup>31</sup> *Ibid.*

<sup>32</sup> Compare Aristotle when he insists that trying to discuss anything at all with someone who seriously questioned whether or not they should be rational would be “like trying to argue with a vegetable.” See Aristotle, *Metaphysics*, trans. by Hugh Lawson-Tancred (Harmondsworth: Penguin, 1998), IV 1006a, 15.

<sup>33</sup> “The relevant sense (of ‘acting’ in which we are always acting) is the sense in which we can always appropriately ask of a conscious human being ‘Why are you doing that?’ Someone who is sitting in a chair staring into space or at a television screen is in that sense acting. We can ask ‘Why have you been sitting in that armchair for the past hour?’ and the answer might be, perhaps, ‘I need to relax’, or ‘I’m watching an interesting programme’, or ‘I’m planning my next lecture.’” Norman, “Practical Reasons and the Redundancy of Motives,” 8.

<sup>34</sup> *Ibid.*, my emphasis.

<sup>35</sup> *Ibid.*, my emphasis.

course, we can be mistaken in taking it that there are reasons to perform a particular action: but exactly the same is the case of taking it that there are reasons to believe something or other. Moral cognitivism does not rule out our making moral mistakes, any more than the view that mathematics is a rational system (call it mathematical cognitivism if you like) rules out our making mathematical mistakes.

There is no need of any theory of motivation at all: for now, there is no longer any mysterious gap between our reasons for doing something and what we do that needs to be bridged by motivation, volition, or whatever. The so-called motives invoked to bridge the gap between reasons and actions, because reasons alone seemed unable to make the connection, turn out not to be needed at all. For instance, it is not that my “failure of moral motivation ... necessarily *involves* or *signals* a cognitive failure”;<sup>36</sup> rather, my failure to act morally well *is* a cognitive failure. In the normal, case we can say all there needs to be said by way of explaining why people do the things they do without referring to motives at all. In fact, motives turn out even metaphorically to feature “only in special cases”<sup>37</sup> where there is some doubt about whether or not the reasons cited by someone are *really* the reasons for which they acted. If “I ask what your *motives* were, I imply that things are not what they seem, that you had *ulterior* motives.”<sup>38</sup> But ulterior motives are just unacknowledged reasons: the question, “What are your (real) motives?” can always be rendered without loss or addition as the question, “What are your (real) reasons?” And of course, motives might be invoked in everyday talk, as they often are, to refer to a person’s general reasons for doing the things they do: “[I]f someone donates money to charity and says that her reason for doing so is the terrible suffering of the victims of the famine, we can say that, if that is her reason, she was motivated by pity.”<sup>39</sup> But to say that is simply to say that pity was her reason for what she did. Even in such unexceptionable everyday talk of motivation, unencumbered by any “theory of motivation,” motives are not something that motivates. They are simply reasons.

## 5. Conclusion: reasons for actions

We do things for reasons. That is just what *doing* something is. And as in the ordinary everyday practical contexts of doing something, so in the specifically moral case: the conviction that moral judgements require something “more” if they are to lead to action is misplaced. To speak out

---

<sup>36</sup> Margaret Olivia Little, “Virtue as Knowledge: Objections from the Philosophy of Mind,” in *Noûs* 31 (1997), 59–79, p. 72.

<sup>37</sup> Norman, “Practical Reasons and the Redundancy of Motives,” 6.

<sup>38</sup> *Ibid.*, 6–7.

<sup>39</sup> *Ibid.*, 7.

against the latest plans to commodify the university, or health provision, because they destroy the requisite co-operative ethos needs no more than the judgement that that is what they will do and that it is wrong. Agency, one might say, goes all the way down. We are a species, individual and collective members of which are fundamentally characterized by agency; to be a person is inescapably to be an agent. And it follows from this that recognizing a consideration as a reason to do something is already to be committed to act on it; again, as Norman insists, “we do not need any further explanation of why, in the normal case, human agents act on their beliefs about what they have good reasons to do. They just do it.”<sup>40</sup> And if that is the case, then, we can rid ourselves of a good deal of entirely unnecessary and confusing philosophical baggage. First, the internal reason/external reason distinction becomes immediately irrelevant: for the point about citing reasons to act in a particular way is precisely “not, then, to show that a belief about right and wrong can explain an action, but that a belief of such a kind can give the agent reason to do or not to do it”—so that “someone who does what is wrong thereby acts in a way that is contrary to reason.”<sup>41</sup> Second, the distinction between motivating and justifying reasons also becomes immediately redundant. Remember, the problem about reasons leading to actions arose because we seemed to need *both* explanation and justification of what people do. To explain is one thing; to justify quite another. True. But the thought is not dependent on any “motivation.” If x justifies z, then x also explains z.

While we are thinking of philosophical baggage and its inconvenience, it is worth pausing to consider how and why some philosophers—noticing that explanation is one thing and justification another, but not wanting to deprive reasons of any motivational role, so as to retain the possibility of moral cognitivism—have tried to square the circle by distinguishing two sorts of reasons: those that actually motivate someone to action, and those that either justify or would justify their action.<sup>42</sup> For people clearly make mistakes about reasons and thus, do things which they suppose there are reasons to do when in fact there are not: so we can explain what they

---

<sup>40</sup> *Ibid.*, 11.

<sup>41</sup> Philippa Foot, *Natural Goodness* (Oxford: Clarendon Press, 2001), 62. As Warren Quinn observes in his discussion of Williams’ “Internal and External Reasons”: “modern subjectivists have extended Hume’s idea that morality produces motives only through its noncognitive content to the idea that it produces reasons only in the same way.” See “Putting Rationality in its Place,” reprinted in W. Quinn’s *Morality and Action* (Cambridge: Cambridge University Press, 1993), 231.

<sup>42</sup> The view once advanced by E.J. Bond in his *Reason and Value* (Cambridge: Cambridge University Press, 1983) that one should distinguish “motivating” from “justifying” reasons was historically a helpful advance on the Humean position; but in the end such a view merely shifts the problem to that of trying to account for their relations, in a way that parallels the “internalism/externalism” debate. If something does not justify a belief or an action—at least to some extent, for there may of course be countervailing reasons—then, it is not any sort of reason.

did without condoning it (she thought she was acting for the best). And with that distinction in place, to state someone's "justifying reason" for what they did (she thought it was cowardly not to speak out) is not to explain it: something else, a motivating reason, is needed for that. Without knowing what "pushed" her to speak, her speaking remains unexplained—unexplained (to go back to the earlier discussion) because its cause has not been identified. Reasons, now "justifying reasons," can't explain actions because they are not affectively engaged. It takes "motivating" reasons to do that. If motivating reasons and justifying reasons are different, as they must be; and if in the case of moral acts they nevertheless have to coincide, as they do if morality is not to be optional, but rather, in Kant's terminology, categorical, then, motivating reasons for a moral act have to take precedence over justifying reasons. And so, we find ourselves with another version of the internal/external problem. But without any "motivational gap" in place, there is no difficulty in seeing that if a judgement justifies an action, then it also explains it. I put on my mac because it is raining. What further explanation is needed for what I did? There are not two kinds of reason, "justifying reasons" and "motivating reasons," but rather two kinds of explanation: intentional and causal. The first explains actions by citing reasons for it; the second explains events by detailing their cause(s).<sup>43</sup>

But still, don't we often do something simply because we want to? And shouldn't that lead us rightly to reassert the view that desires sometimes motivate at least some actions, so that we need the language of motivation after all? No. Very often, of course, we do say that we are going to do something "because we want to": but the "because" here only gives an appearance of anything causal going on. Rather, our wanting to do something *is* the reason for doing it; and it is sometimes a perfectly good reason. But wanting to do something does not cause us to do it; it does not compel us to act. Otherwise, again, we would not be *acting* at all, but merely behaving—as indeed we sometimes do, when we just respond to some stimulus such as a desire. But automatically, having another glass of wine, rather than deciding to have another, or simply picking something off the supermarket shelf, rather than choosing it, are not *actions* at all.

It is also the case, of course, that we do sometimes talk of reasons as motivating us to adopt certain beliefs. You get yourself to (try to) believe that the job you are doing is not one that exploits people in order to be able to get on with doing it without too much trouble. Or, you give assent to a judgement, and succeed in making it your own, in order to join, or to remain

---

<sup>43</sup> What exactly the relation might be between these constitutes one version of the problem of free will and determinism. That, however, is another matter. But if determinism were true, then of course, we would not—could not—be agents.



within, some group or another: consider Members of Parliament and the rubbish their parties require them to believe if they are to have any hope of preferment. Nor is this always a matter of simply appearing to have the right beliefs and to make the right judgements: cynicism is, at least, sometimes matched by sincerity, which is what makes the histories of many politicians' changing beliefs and judgements so puzzling. How can she sincerely say today what yesterday she denied? How can he now judge that something is indeed the case which yesterday he was clearly convinced was not? But again, the terminology of reasons as motivating us to adopt beliefs is entirely unnecessary; and in terms of understanding actions, it is profoundly misleading. There is no difference between saying that a politician was motivated to adopt some belief in the hope of preferment, and saying that their hope of preferment was the reason for their adopting the belief. Where people are said to "have motives," what is actually being claimed is that there are reasons for what they do or intend to do.

Of course, there being a reason for you or me to do something does not always result in our doing it. But again, that is not because of some sort of motivational failure. We might not be aware of the reason; we might be aware of it, but underestimate its importance relative to countervailing reasons; or we might just feel too tired or too ill to do this particular thing, or even to do anything at all. It is a rational failure: one that is no less, but also no more, mysterious than rational failures regarding beliefs. Of course, there are ways in which rational failure is indeed puzzling, hence, the invocation of "weakness of will," for instance, or anomie.<sup>44</sup> But its ubiquity in the context of action no more undermines the rationality of morality than does its ubiquity in the context of beliefs. Our mistakes in reasoning about the world or about mathematics do not undermine the rationality of such reasoning. Why should mistakes in practical reasoning—including moral reasoning—be thought to undermine the rationality of such reasoning? It is only because the action concerned is conceptualized as some sort of "addition" to the judgement that a problem appears to arise. But on a view of agency as the "standard" state of affairs, there is no more a question of judgement plus action than there is of judgement plus belief. In each case, these two elements are part of a single whole.

*Centre for Applied Philosophy, Politics, & Ethics,  
University of Brighton, United Kingdom*

---

<sup>44</sup> A defence of moral cognitivism would, of course, have to deal in detail with these and other matters. Here, however, it is enough to dispel the view that the phenomenon of motivation offers a threat to moral cognitivism.

## References

- Aristotle, *Metaphysics*, trans. by Hugh Lawson-Tancred (Harmondsworth: Penguin, 1998).
- Brecher, Bob, *Getting What You Want? A Critique of Liberal Morality* (London & New York: Routledge, 1997).
- Bond, E.J., *Reason and Values* (Cambridge: Cambridge University Press, 1983).
- Dancy, Jonathan, "Why There is Really No Such Thing as the Theory of Motivation," in *Proceedings of the Aristotelian Society*, 95 (1995).
- Davidson, Donald, "Actions, Reasons, and Causes," reprinted in his *Essays on Actions and Events* (Oxford: Oxford University Press, 1980).
- Foot, Philippa, *Natural Goodness* (Oxford: Clarendon Press, 2001).
- Hampton, Jean, *The Authority of Reason* (Cambridge: Cambridge University Press, 1998).
- Hobbes, Thomas, *Leviathan* (Harmondsworth: Penguin, 1968).
- Hollis, Martin, *The Cunning of Reason* (Cambridge: Cambridge University Press, 1987).
- Kant, Immanuel, *Groundwork of the Metaphysics of Morals*, trans. Mary Gregor (Cambridge: Cambridge University Press, 1997).
- Little, Margaret Olivia, "Virtue as Knowledge: Objections from the Philosophy of Mind," in *Noûs*, 31 (1997).
- Norman, Richard, "Practical Reasons and the Redundancy of Motives," in *Ethical Theory and Moral Practice*, 4 (2002).
- \_\_\_\_\_, "Public Reasons and the 'Private Language' Argument," in *Philosophical Investigations*, 23 (2000).
- Quinn, Warren, "Putting Rationality in its Place," reprinted in his *Morality and Action* (Cambridge: Cambridge University Press, 1993).
- Sayre-McCord, Geoffrey, "Coherentist Epistemology and Moral Theory," in *Moral Knowledge?*, ed. by W. Sinnott-Armstrong and M. Timmons, (Oxford: Oxford University Press, 1996).
- Williams, Bernard, "Replies," in *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*, ed. by J.E.J Altham and Ross Harrison, (Cambridge: Cambridge University Press, 1995).

## Where Epistemic Safety Fails

---

**Mark Anthony L. Dacela**

**Abstract:** In a previous paper, I briefly profiled unsafe beliefs as either: (1) beliefs formed using a method that is conditionally reliable and (2) beliefs formed using a method with unstable reliability. I dubbed these profiles as B-type and C-type, respectively. Extending this analysis, I will demonstrate how these belief types operate and why they fail in some notable counterexamples to safety offered by Neta and Rohrbaugh, Cosmesaña, Baumann, Kelp, Bogardus, and Freitag. Examining these cases also motivate my thesis that a method's conditional reliability or instability does not render a belief formed by an actually reliable method unjustified; its epistemic worth remains intact, unsafe as it may be.

**Keywords:** epistemology, safety, knowledge, possible worlds

### Introduction

In a previous paper, I briefly profiled unsafe beliefs as either: (1) beliefs formed using a method that is conditionally reliable and (2) beliefs formed using a method with unstable reliability.<sup>1</sup> I dubbed these profiles as B-type and C-type, respectively. Extending this analysis, I will demonstrate how these belief types operate and why they fail in some notable counterexamples to safety offered by Neta and Rohrbaugh,<sup>2</sup> Cosmesaña,<sup>3</sup>

---

<sup>1</sup> See Mark Anthony L. Dacela, "Are Modal Conditions Necessary for Knowledge?" in *Kritike: An Online Journal of Philosophy*, 13:1 (2019), 101–121.

<sup>2</sup> See Ram Neta and Guy Rohrbaugh, "Luminosity and the Safety of Knowledge," in *Pacific Philosophical Quarterly*, 85 (2004), 396–406.

<sup>3</sup> See Juan Cosmesaña, "Unsafe Knowledge," in *Synthese*, 146 (2005), 395–404.

Baumann,<sup>4</sup> Kelp,<sup>5</sup> Bogardus,<sup>6</sup> and Freitag.<sup>7</sup> Examining these cases also motivate my thesis that a method's conditional reliability or instability does not render a belief formed by an actually reliable method unjustified; its epistemic worth remains intact, unsafe as it may be.

This paper is divided into the following sections: first, a quick review of the safety condition and its use of possible worlds; then, a discussion of my profiling of *unsafe beliefs*; and finally, an analysis of *unsafe beliefs* in the counterexamples cited above.

### Safety in Brief

Sosa offered safety as a necessary condition for knowing.<sup>8</sup> He stated this condition, where "S" stands for subject and "p" stands for proposition, as:

S's belief is safe = df. S would believe that *p* if it were so that *p* or alternatively S would not believe that *p* without it being the case that *p*.

We may simplify this condition using this subjunctive conditional:

If S were to believe *p*, it would be the case that *p*.

And employing the possible-worlds account of subjunctive conditionals, we can modify this to:

S's belief is safe = df. In the closest possible worlds in which S believes *p*, *p* is true.

Sosa's analysis of knowledge can then be expressed as:

---

<sup>4</sup> See Peter Baumann, "Is Knowledge Safe?" in *American Philosophical Quarterly*, 45:1 (2008), 19–30.

<sup>5</sup> See Christoph Kelp, "Knowledge and Safety," in *Journal of Philosophical Research*, 34 (2009), 21–31.

<sup>6</sup> See Tomas Bogardus, "Knowledge under Threat," in *Philosophy and Phenomenological Research*, 88:2 (2014), 289–313.

<sup>7</sup> See Wolfgang Freitag, "Safety, Sensitivity and 'Distant' Epistemic Luck," in *Theoria* 80:1 (2014), 44–61.

<sup>8</sup> See Ernest Sosa, "How to defeat Opposition to Moore," in *Philosophical Perspectives* 13 (1999), 141–153.

S knows that  $P = \text{Df.}$  (1)  $p$  is true, (2) S believes that  $p$ , (3) in the closest possible worlds in which S believes  $p$ ,  $p$  is true (safety condition).

To understand how safety works, we need to review the semantics at play. From here on I will use the terms “subjunctive conditional” and “subjunctive” interchangeably. Also, note that moving forward “‘ $pq$ ’” represents the subjunctive: *If it were  $p$  then it would have been  $q$ .* Now a brief note on to Lewis’s and Stalnaker’s accounts of subjunctives: these theories were offered as ways of determining the truth condition of subjunctives. The question that these theories try to answer can then be stated as, “When do we judge statements in the form ‘ $pq$ ’ as true?”

Consider first Stalnaker’s account. (Let “@” stand for actual world, and, “ $p$ -world” for world where the antecedent is true):

**STL:** ‘ $pq$ ’ is true in @ = Df. ‘ $pq$ ’ is true in the closest  $p$ -world to @.

Stalnaker asks us to consider the world closest to the actual world, which for him refers to the world which “differs minimally” from the actual world, and in which the antecedent ( $p$ ) is true: If ‘ $pq$ ’ is true in that world then ‘ $pq$ ’ is true in @.<sup>9</sup> So, given a set of  $p$ -worlds we only check the  $p$ -world which differs minimally to @. Stalnaker also tells us that requiring a world that “differs minimally” implies that:

[T]here are no differences between the actual world and the selected world except those that are required, implicitly or explicitly, by the antecedent . . . [and] among the alternative ways of making the required changes, one must choose one that does the least violence to the correct description and explanation of the actual world.<sup>10</sup>

These further conditions recognize that fact that different situations may obtain in worlds where the antecedent of a given subjunctive is true. And that there is a *degree of variance*, such that one world is more similar to the base world (i.e., the actual world) than another world. We check only the  $p$ -world

---

<sup>9</sup> Robert Stalnaker, “A Theory of Conditionals,” in *Studies in Logical Theory*, ed. by Nicholas Rescher (Oxford: Basil Blackwell, 1968), 102.

<sup>10</sup> *Ibid.*, 104.

that is most similar in that it “differs minimally” from the actual world. Lewis calls this the Stalnaker assumption:<sup>11</sup>

For every world @ and antecedent  $p$  that is accessible to @, there is a sphere around @ containing exactly one  $p$ -world.

Lewis rejects this assumption and offers this revised account:<sup>12</sup>

**LEW:** ‘ $pq$ ’ is true in @ = Df. Some world in which  $p$  and  $q$  are true is closer to @ than any world in which  $p$  and not- $q$  is true.

Again, *closeness* refers to the similarity relation of worlds. Unlike Stalnaker, Lewis does not limit the set of *relevant worlds* to only one member. Lewis’s account asks us to compare worlds in which ‘ $p$  and  $q$ ’ obtains and worlds in which ‘ $p$  and not- $q$ ’ obtains. If at least one member of the first set is more similar to the base world (or the actual world) than any member of the second set, then the subjunctive ‘ $pq$ ’ is true.

By adding a temporal element to the equation, Lewis further qualifies the set of relevant worlds, where “ $w_1$ ” stands for a possible world:<sup>13</sup>

**LEW\*:** ‘ $pq$ ’ is true at @, where  $p$  is entirely about affairs in a stretch of time  $t_1$  = Df. (1)  $p$  is true at  $w_1$ ; (2)  $w_1$  is exactly like @ at all times before the transition period beginning shortly before  $t_1$ ; (3)  $w_1$  conforms to the actual laws of nature at all times after  $t_1$ ; and (4) during  $t_1$  and the preceding transition period,  $w_1$  differs no more from @ than it must to permit  $p$ .

LEW\* tells us just how similar the worlds should be. (1) limits the relevant set to  $p$ -worlds (same as Stalnaker), which we take as the initial similarity test. (2) to (4) set a *similarity range*: in (2) the worlds should be exactly similar from all times before the transition period which begins shortly before  $p$  obtains ( $t_1$ ); in (3) the laws of nature should be similar at all times; and in (4) and onwards, the difference should be no more than what it is required for  $p$  to obtain.

---

<sup>11</sup> David Lewis, *Counterfactuals* (Oxford: Blackwell Publishers, 1973), 78.

<sup>12</sup> *Ibid.*, 82.

<sup>13</sup> David Lewis, “Counterfactual Dependence and Time’s Arrow,” in *Noûs* 13:4 (1979), 462.



Lewis also offers a priority list in weighing factors for similarity (in order of significance): (1) avoid big, widespread, diverse violations of law (large miracles); (2) maximize the spatiotemporal region throughout which a perfect match of particular facts prevails; (3) avoid even small, localized, simple violations of law (small miracles); and (4) secure approximate similarity of particular facts.<sup>14</sup> (1) to (4) tell us that a perfect match of small facts for an extended time counts more than the absence of small miracles in weighing overall similarity. However, the absence of large miracles outweighs these two factors.

Going back to safety, note that this condition is in the form of a counterfactual. And as discussed, to check if a counterfactual is true, we need to check possible worlds in which the antecedent is true and see if the consequent holds there as well. Safety thus requires us to check close possible worlds where the subject believes the proposition and see if in those worlds the proposition is true. Then alternatively, in the close possible worlds where the subject does not believe the proposition, the proposition is false.

### Unsafe Beliefs

I offered two profiles of unsafe beliefs: B-type and C-type.<sup>15</sup> B-type beliefs are formed using a conditionally reliable method, while C-type beliefs are formed with unstable reliability. Developing the notion introduced by Goldman,<sup>16</sup> we can consider a method conditionally reliable if in case there is a possible circumstance in which it fails to produce a true belief; and unstable if at any given instance in can produce a false belief. We can say then that methods with unstable reliability are also conditionally reliable, but not all conditionally reliable methods are unstable.

To appreciate the difference, it is helpful to think of this in modal terms. A method is conditionally reliable if there are worlds in which it produces false beliefs. If these worlds are extremely close to the actual world, the method is unstable. We can also think of it in terms of probability. If the probability that the method will produce a false belief is high, the method is unstable. If there is a possibility that it will produce a false belief but the probability is low, then the method is conditionally reliable.

I also identified other features of both B-type and C-type beliefs. First, both beliefs are internally justified. This means that what justifies the belief is within the conscious grasp of the subject. In other words, the evidence is

---

<sup>14</sup> *Ibid.*, 472.

<sup>15</sup> Dacela, "Are Modal Conditions Necessary for Knowledge?" 104.

<sup>16</sup> See Alvin I. Goldman, "What is Justified Belief?" in *Justification and Knowledge*, Philosophical Studies Series in Philosophy, Vol. 17, ed. by George Sotiros Pappas (Dordrecht: Springer, 1979), 1–23.

known. Second, both beliefs are factually defeated. Following Steup, a factual defeater is a true proposition, hidden from the subject, and either weakens the justification of a belief or renders it completely unjustified.<sup>17</sup>

### Counterexamples to Safety

Now we take a closer look at some notable counterexamples to safety to see how these beliefs operate and why they fail to meet the safety requirement. As a way of framing my analysis note that at least four sets of possible (epistemic) worlds are at play in these cases: {} worlds in which the proposition is true and the subject believes it, {} worlds in which the subject falsely believes the proposition, {} worlds in which the proposition is true but the subject does not believe it, and {} worlds in which the proposition is false and the subject does not believe it:

{} Bsp. p  
 {} Bsp. p  
 {} Bsp. p  
 {} -Bsp, p

In each case, the crucial step is determining if these sets are included or excluded in the set of relevant or close worlds {}. The similarity criterion states that any member of {} is similar to the actual world (@):

For a given world, call it the actual world @, and a possible world #, # is a member of {} iff # is similar to @.

Note that a belief is safe if and only if ‘*if it were that the subject believes the proposition, the proposition is true*’. The safety condition then limits the set of relevant or close worlds to {} worlds in which the proposition is true and the subject believes it; and excludes {} worlds in which the subject falsely believes the proposition:

**Belief is safe iff:** (3) {} includes members of {} and  
 (4) {} excludes members of {}.

(3) and (4) are necessary conditions.

---

<sup>17</sup> Matthias, Steup, *An Introduction to Contemporary Epistemology* (Upper Saddle River, NJ: Prentice-Hall, 1996), 14.

*Gottit and Nogood (Baumann)*

Baumann presents a case that exposes the safety condition's lack of clarity and straightforwardness.<sup>18</sup> Consider this version first:

**MASK.** The following story is from Milleville, a small town in the Wild West. Two notorious bank robbers have been doing business in the area for some time: Frederick P. Nogood and Wilbur Gottit. Their faces are on "Wanted" posters all over the place. They are rivals and don't like each other at all. When Nogood goes to the bank, he uses a perfectly deceptive Gottit mask; when Gottit goes to the bank, he uses a perfectly deceptive Nogood mask. Nobody but they themselves know this. One day, Frank is walking around in the streets of Milleville when he suddenly sees a bank robber leave the bank with a bag full of money on his back, shooting back at the bank. Frank happens to look at him and there is no doubt for him: It is Nogood. But it really is Gottit with his Nogood mask on. However, by sheer coincidence Gottit's Nogood mask slips at that very moment, and Frank notices all this. This is extraordinary because something like that only happens this one time to Gottit and never to Nogood.

So, Frank forms the belief that *Gottit is the robber* ( $p$ ). And clearly, Frank knows  $p$ . However, Baumann claims that Frank's belief does not satisfy the safety condition, since there are close worlds where he falsely believes  $p$ . In these worlds, Frank believes that Gottit just robbed a bank when it was really Nogood wearing his Gottit mask. The counterfactual 'S would believe that  $p$  only if it were so that  $p'$  does not hold in this case.

Now a safety theorist might question just how close the world where Nogood is wearing his Gottit mask ( $w_1$ ) to the actual world where Frank notices Gottit's Nogood mask slip (@). She might say that only worlds where Frank sees Gottit's Nogood mask slip should be counted as close worlds. That is, worlds where everything is the same with the actual world except for one small epistemically irrelevant detail (e.g., Gottit has one less hair on his right leg), or something slightly different happening far elsewhere that does not have anything to do with Frank, Gottit, or Nogood. But Baumann questions just how defensible this notion of closeness would be.

---

<sup>18</sup> Baumann, "Is Knowledge Safe?" 20.

Baumann identifies and evaluates possible determinants of closeness. He started with this general condition, where @ stands for the actual world and  $w_1$  for the possible world in question:

**D1**  $w_1$  is close to @ = Df. The differences between  $w_1$  and @ are epistemically irrelevant (enough).

D1 recognizes that some differences are *epistemically irrelevant* while others are not. Whether or not the difference is relevant depends on how much it varies the *epistemic situation* of the subject in @. That Frank has one less hair on his right leg in  $w_1$  does not change the epistemic situation. So we consider this difference epistemically irrelevant. Thus, we can modify D1 to:

**D1\***  $w_2$  is close to @ = Df. The epistemic situation of S is the same in both  $w_1$  and  $w_a$ .

Baumann still finds D1\* unsatisfactory, since it does not tell us what an epistemic situation is, and, more importantly, what makes an epistemic situation the same or different. He also finds other versions problematic:<sup>19</sup>

S's epistemic situation is the same in  $w_1$  and  $w_a$  =

**D2** Df. S holds the same belief in  $w_1$  and @.

**D3** Df. The truth value of S's belief is the same in  $w_1$  and @.

**D4** Df. The relevant facts are the same in both  $w_1$  and @.

**D5** Df. The initial conditions are the same in both  $w_1$  and @.

Baumann claims that they are either trivially true (D2), or they make safety trivial (D3 to D5). He then closely examines three promising versions. Consider this first:

**D6** S's epistemic situation is the same in  $w_1$  and @ = Df. S's warrant for believing  $p$  is the same in  $w_1$  and @.

Baumann notes that *warrant* here is taken in its broadest sense.<sup>20</sup> It includes the *reasons or justification* the subject might have for the belief and the *methods of belief acquisition*. D7 involves sameness of *subjective* evidence. It asks us to

---

<sup>19</sup> *Ibid.*, 23.

<sup>20</sup> *Ibid.*

## 62 WHERE EPISTEMIC SAFETY FAILS

consider the way the evidence appears to the subject. We can restate it this way:

**D6\*** S's epistemic situation is the same in  $w_1$  and @ = Df.  
(1) Subjectively speaking, S has the same evidence for her belief in  $w_1$  and @.

Now consider these two worlds:

@ Frank notices Gottit's Nogood mask slip ( $e_a$ ),  
then forms the belief *that Gottit is the robber* ( $p$ ).  
 $w_1$  Frank sees Nogood wearing his Gottit mask ( $e_1$ ),  
then forms the belief *that Gottit is the robber* ( $p$ ).

Note that as far *seeing* goes, Frank's evidence in both worlds is the same: Gottit's face. Given D6 then, Frank has the same epistemic situation in both worlds. D6 does not restrict the set of close worlds to worlds where Gottit's mask slip. So, D6 does not work if the idea is to exclude worlds like  $w_1$ .

Now consider this definition:

**D7** S's epistemic situation is the same in  $w_1$  and @ = Df.  
Objectively speaking, S has the same evidence for believing  $p$  in  $w_1$  and @.

Given D7, the safety theorist can claim that Frank's epistemic situation in  $w_1$  and @ vary. In @, Frank actually sees Gottit's face. While in  $w_1$ , he is in fact seeing Nogood in his Gottit's mask. Frank may not be able to tell the difference, but objectively speaking, his evidence in  $w_1$  is different from his evidence in @. However, Baumann finds D8 too strong and not very illuminating.<sup>21</sup> Typically, the subject's evidence for his belief  $p$  in a world where  $p$  is true would differ from his evidence for the same belief in a world where  $p$  is false. This is the case with worlds @ and  $w_1$ . In @, Frank's belief that *Gottit is the robber* ( $p$ ) is true, and his evidence confirms this, while in  $w_1$ , Frank's belief is false but his evidence misleads him to believe otherwise. D7 thus excludes worlds in which the subject's belief is false. Baumann worries that this would trivialize the safety account.<sup>22</sup>

Finally, consider this definition:

---

<sup>21</sup> *Ibid.*, 24.

<sup>22</sup> *Ibid.*

D8S's epistemic situation is the same in  $w_1$  and @ = Df. S's belief forming method is the same in  $w_1$  and @.

D8 does not work with the set of worlds we have. It does not tell us how Frank's method in @ is relevantly different from his method in  $w_1$ . Baumann thinks D8 run into similar problems in the argument from sameness and differences of evidence or reasons (see discussion above). On the one hand, if you consider *seeing* or *perceiving* as Frank's belief forming method in @, then the difference in method does not seem relevant. In  $w_1$  Frank's belief is formed via *perception* as well (only he's actually seeing Nogood's Gottit mask). On the other hand, if we construe method in the externalist sense, then they only differ in terms of the truth-value of the proposition. Everything else would be the same (Frank uses the method of looking at the person's face in both worlds) except that in the @ the belief is true, and in  $w_1$ , false. This leads to the exclusion of worlds where the subject's belief is false, which threatens to trivialize safety.

Now consider version two of Baumann's case:<sup>23</sup>

**FAKE.** Many people do robberies in the Milleville area. All of them (including Nogood) wear non-slipping perfect Gottit masks, except Gottit who usually wears a Nogood mask, except today. Frank happens to see Gottit without his mask (he forgot to bring it to work today).

So, Frank forms the belief that *Gottit is the robber* ( $p$ ). Baumann thinks it is uncontroversial to claim in this case that Frank does not know  $p$ . And this is consistent with safety. There are close worlds where Frank *falsely* believes  $p$ . The subjunctive conditions 'S would believe  $p$  only if it were so that  $p'$  does not hold in this case. Frank's belief is not safe. Notice that those worlds where another person wears a Gottit mask are considered close in this case. While in *Mask*, the world where Nogood wears a Gottit mask, arguably, is not included in the set of close worlds. Baumann wonders why this is so. *Mask* and *Fake* differ in two ways: (1) There are more robbers in *Fake* not just Gottit and Nogood and (2) Gottit wears a slipping mask in *Mask* but not in *Fake*.<sup>24</sup> Bauman argues that neither of these explains why the set of close worlds or the *ceteris paribus* set varies in *Mask* and *Fake*.

Baumann considers (1) negligible.<sup>25</sup> You can easily modify *Mask* to include many masked robbers. This would not significantly change the result. Frank still knows that *Gottit is the robber* ( $p$ ) yet Frank's belief remains *unsafe*.

---

<sup>23</sup> *Ibid.*, 25.

<sup>24</sup> *Ibid.*

<sup>25</sup> *Ibid.*

(2) does not solve the puzzle either. It would explain why the *ceteris paribus set* varies in *Mask* and *Fake* only if it would imply a difference in either reasons or methods. But even if you grant these differences, it is not clear why such qualitative differences would have implications on the *ceteris paribus set* (this argument would parallel the ones discussed above). Without an argument to explain why these differences relevantly vary the *ceteris paribus set*, safety theorists should assume that in both *Mask* and *Fake* either: the masked worlds are included or excluded in the *ceteris paribus set*. Baumann asserts that either way the counterexample would hold.

### *Halloween Party (Cosmesaña)*

Now consider Juan Cosmesaña's example:

**HALLOWEEN:** There is a Halloween party at Andy's house, and I am invited. Andy's house is very difficult to find, so he hires Judy to stand at a crossroads and direct people towards the house (Judy's job is to tell people that the party is at the house down the left road). Unbeknownst to me, Andy doesn't want Michael to go to the party, so he also tells Judy that if she sees Michael she should tell him the same thing she tells everybody else (that the party is at the house down the left road), but she should immediately phone Andy so that the party can be moved to Adam's house, which is down the right road. I seriously consider disguising myself as Michael, but at the last moment I don't. When I get to the crossroads, I ask Judy where the party is, and she tells me that it is down the left road.<sup>26</sup>

In *Halloween*, Cosmesaña knows that *the party is down the left road* (*p*), but he would have believed this even if it weren't true. The subjunctive conditional 'S would believe that *p* only if it were so that *p'* does not hold in this case. Cosmesaña's belief is not safe, yet he knows *p*.

But how do we motivate the intuition that Cosmesaña knows *p* in this case? One way is to point out that his basis for belief *p*, Judy's testimony (*t*), is at least *actually* reliable, although it *possibly* is not. A method or a belief-basis is reliable if it is knowledge conducive. Basis *t* would have been unreliable in a possible world where he is disguised as Michael. Call this *possible unreliability* to distinguish it from *actual* reliability. In this possible

---

<sup>26</sup> Cosmesaña, "Unsafe Knowledge," 397.



world, he would have falsely believed  $p$  on the same basis,  $t$ . This makes his belief unsafe. So, if Cosmesaña considered disguising himself as Michael,  $t$  would have been questionable. But this is not the case. So,  $t$  is *actually* reliable. This makes Cosmesaña's belief justified. And that warrants the intuition that he *knows*  $p$ .

Cosmesaña's example demonstrates that knowledge tolerates this sort of weak reliability where a belief-basis would have easily been unreliable though it actually is not. In this case, safety requires that my basis for believing  $p$ ,  $t$ , produce a true belief in all the close possible worlds (where I use  $t$  in forming belief  $p$ ). But Cosmesaña claims that this is not necessary for knowledge.<sup>27</sup>

### *Russell's Clock (Kelp)*

Now here is a case that involves a *counterfactual intervener*, offered by Christoph Kelp:

**CLOCK.** Suppose Russell's arch-nemesis has an interest that Russell forms a belief (no matter whether true or not) that it's 8:22 by looking at the grandfather clock when he comes down the stairs. Russell's arch-nemesis is prepared to do whatever it may take in order to ensure that Russell acquires a belief that it's 8:22 by looking at the grandfather clock when he comes down the stairs. (Since we are concerned with a conceptual claim here, Russell's arch-nemesis may have means available to do so that we can imagine only in our wildest dreams. For instance, he may be an evil-demon who can set the clock to 8:22 with his invisible hand a second before Russell looks at it.) However, Russell's arch-nemesis is also lazy. He will act only if Russell does not come down the stairs at 8:22 of his own accord. Suppose, as it so happens, Russell does come down the stairs at 8:22. Russell's arch-nemesis remains inactive. Russell forms a belief that it's 8:22 ( $p$ ). It is 8:22. The grandfather clock is working reliably as always.<sup>28</sup>

Kelp claims that Russell knows it is 8:22 ( $p$ ) since,

---

<sup>27</sup> *Ibid.*, 402.

<sup>28</sup> Kelp, "Knowledge and Safety," 27–28.

[H]e looks at a perfectly working clock, he has the ability to read the clock, exercises his ability and hits upon the truth through the exercise of this ability. Moreover, his belief is true. It is in fact 8:22.

But he also points out that Russell's belief is not safe since at all the nearby worlds he would have falsely believed  $p$ .<sup>29</sup> These are worlds where Russell comes down a minute earlier or later prompting his arch-nemesis to intervene and change the clock's setting to 8:22. In these possible worlds, Russell would have still believed that it's 8:22 ( $p$ ), even if it weren't. The subjunctive "S would have believed  $p$  only if it were so that  $p$ " does not hold in this case. And yet, Russell knows  $p$ .

There are striking similarities between *Halloween* and *Clock*. In both examples, the subject's basis for forming the belief in question has a weak sort of reliability similar to Cosmesaña's belief-basis in *Halloween*. That it is actually reliable motivates the intuition that the subject knows. That it would have been unreliable makes the belief unsafe. Russell's grandfather clock is *actually* reliable since his arch-nemesis did not intervene in the actual world, though he would have in close possible worlds where Russell comes down earlier or later. This latter bit makes the grandfather clock possibly or potentially unreliable.

Kelp however claims that his counterexample is more plausible than Cosmesaña's; in fact, the latter's argument strikes him as unconvincing.<sup>30</sup> He points out that in *Halloween*, a lot of things had to be different for the subject to have a false belief: Cosmesaña has to decide to disguised himself as Michael, he must have successfully done so, Judy must be convinced that he is Michael, she must have phoned Andy, Andy must have moved the party elsewhere. While in *Clock*, Kelp argues that all it takes is that Russell stays in bed a minute longer or he comes down a minute earlier.<sup>31</sup>

But is Kelp correct in claiming this? Consider these three worlds:

$w_1$  Russell comes down at 8:21, looks at the clock then forms the belief that it's 8:22 ( $p$ ).

@ Russell comes down at 8:22, looks at the clock then forms the belief that it's 8:22 ( $p$ ).

$w_2$  Russell comes down at 8:23, looks at the clock then forms the belief that it's 8:22 ( $p$ ).

---

<sup>29</sup> *Ibid.*, 27.

<sup>30</sup> *Ibid.*, 25.

<sup>31</sup> *Ibid.*, 28.

Kelp is claiming that the only difference between @ and  $w_2$  is that, in  $w_2$ , Russell stays in bed a minute longer. But why would Russell stay in bed a minute longer? Definitely something else would have to change. It could either be something internal, viz., Russell is not as eager or motivated to wake up and start his day, or something external, viz., his alarm clock was set at 8:23. Any of this would imply some other changes. For instance, how would you explain Russell's unwillingness to get off his bed? Perhaps he has a meeting with someone he does not really like. Or there is some chore he has to do that day. This holds true with the other possibility; a lot of things need to vary to explain why Russell set his alarm clock at 8:23 instead of 8:22.

The same can be said about @ and  $w_1$ . Kelp thinks that the only difference between these two worlds is that, in the latter, Russell would have come down a minute earlier.<sup>32</sup> But this would certainly imply other things as well. Perhaps this time Russell is motivated to start his day, or he set his alarm clock at 8:21. And both would imply other changes too. Moreover, in  $w_1$  and  $w_2$ , Russell's arch-nemesis decided to intervene, and have done so successfully. And, both worlds, the clock's hand is pointing at a different number. The point here is that what varies in worlds is never just one small detail. However, without a clear way to determine which worlds are close, Kelp's counterexample would still hold against safety.

### *Water and Flashes (Neta and Rohrbaugh)*

We now turn to two cases, call them *Water* and *Flashes*, presented by Ram Neta and Guy Rohrbaugh:<sup>33</sup>

**WATER.** I am drinking a glass of water which I have just poured from the bottle. Standing next to me is a happy person who has just won the lottery. Had this person lost the lottery, she would have maliciously polluted my water with a tasteless, odorless, colorless toxin. But since she won the lottery, she does no such thing. Nonetheless, she *almost* lost the lottery. Now, I drink the pure, unadulterated water and judge, truly and knowingly, that *I am drinking pure, unadulterated water* (p). But the toxin would not have flavored the water, and so had the toxin gone in, I would still have believed falsely that I was drinking pure, unadulterated water. The actual case and the envisaged possible case are extremely similar in

---

<sup>32</sup> *Ibid.*

<sup>33</sup> Neta and Rohrbaugh, "Luminosity and the Safety of Knowledge," 399–400.

all past and present phenomenological and physical respects, as well as nomologically indistinguishable. (Furthermore, we can stipulate that, in each case, I am killed by a sniper a few moments after drinking the water, and so the cases do not differ in future respects.) Despite the falsity of my belief in the nearby possibility, it seems that, in the actual case, I know that I am drinking pure, unadulterated water.

**FLASHES.** I am participating in a psychological experiment, in which I am to report the number of flashes I recall being shown. Before being shown the stimuli, I consume a glass of liquid at the request of the experimenter. Unbeknownst to either of us, I have been randomly assigned to the control group, and the glass contains ordinary orange juice. Other experimental groups receive juice mixed with one of a variety of chemicals which hinder the functioning of memory without a detectable phenomenological difference. I am shown seven flashes and judge, truly and knowingly, that *I have been shown seven flashes* ( $p$ ). Had I been a member of one of the experimental groups to which I was almost assigned, I would have been shown only six flashes but still believed that I had been shown seven flashes due to the effects of the drug. It seems that in the actual case I know that the number of flashes is seven despite the envisaged possibility of my being wrong. And yet these possibilities are as similar in other respects as they would have to be for the experiment to be well designed and properly executed.

In both cases, I know  $p$ , yet my knowledge is not safe: there is a nearby world in which I *falsely* believe  $p$ .

In *Water*, in the close possible world where the person next to me lost the lottery, she would have spiked my drink with a phenomenologically and physically undetectable toxin. And this would have falsified  $p$ , yet I would have still believed it. Similarly, in *Flashes*, in the close possible world where I am assigned to one of the experimental groups, I would have shown only six flashes. This would have falsified  $p$ , yet due to the effect of the drug given to me, I would have still believed it.

Neta and Rohrbaugh claims that in both cases the possible worlds in which the subject *falsely* believes  $p$  are initially similar in just about every

aspect except for the truth of  $p$  to the *actual* world in which he knows  $p$ .<sup>34</sup> We should note, however, that other changes are at play here. In *Water*, it seems clear that these things would also vary between the actual world and the nearby world: the lottery result, the subject's mood, and as a consequence, the action of the person next to me, and the quality of the water I drank (this falsifies my belief  $p$ ). Notice how these changes are linked. In the close possible world where I falsely believed  $p$ , call these  $W_{2a}$ , the person next to me does not have the winning ticket, and that makes her unhappy (perhaps bitter is more accurate) so much so that she spikes my drink with toxin. Her actions, obviously, compromises the quality of my drink. In *Flashes*, what vary are the grouping assignment, the quality of my drink, the reliability of my memory, and the number of flashes (this falsifies my belief that  $p$ ). These are not isolated changes either. In the close possible world where I falsely believed  $p$ , call these  $W_{2b}$ , I was assigned to a group which members are asked to drink a spiked orange juice, which then compromises the reliability of my memory, making me believe that I saw seven flashes, forgetting that I only saw six.

But are these changes enough to disqualify  $w_{2a}$  and  $w_{2b}$  as nearby worlds to the actual ones, in both cases, where I know  $p$ ? Let's check for nearby possible worlds more similar to the ones considered as actual in both cases. In *Water*, a more similar world to the actual than  $w_{2a}$  is one in which the lottery result is the same. Similar to the actual world, the person next to me wins the lottery. And this leads to a series of events that make my belief  $p$  true. Any other changes would be inconsistent with the realities we have established in describing the actual world, and this too would warrant other changes that would make this possible world significantly different from the actual one. The same goes in *Flashes*. A closer world to the actual one than  $W_{2b}$  is a world where I was assigned to the same group. This also triggers a series of events that eventually make my belief  $p$  true.

Notice then that in these examples, the actual events are closely linked to each other, viz., a slight change in the initial conditions would vary the truth-value of the proposition. The lottery result and the assigning of groups are both crucial, since these events determine what happens next, and, eventually, whether my belief is true or false. If the person next to me wins, I would have truly believed  $p$ . Otherwise I would have been mistaken. If I were assigned to the non-experimental group, I would have truly believed  $p$ . If I were assigned to the other group,  $p$  would have been false. Also, note that these are the conditions set in the actual world described in both cases. In *Water*, it was stated that the person next to me did not poison my drink because she won the lottery. Had this not been the case, I would have falsely

---

<sup>34</sup> *Ibid.*, 399.

believed that *I am drinking pure, unadulterated water* ( $p$ ). While in *Flashes*, it was stated that, if I were assigned to the experimental group, I would have falsely believed that *I saw seven flashes* ( $p$ ).

These conditions will help us identify some non-relevant worlds. For instance, in *Water*, we've established that *she would pollute my drink only if she did not win the lottery*. So we consider non-relevant the possible world in which she pollutes my drink after winning the lottery. The condition we've set in describing the actual world gives us reason to think that this could not have easily been the case. So a world in which this obtains is not a relevant world. In the same way, a world in which she does not pollute my drink after losing the lottery is also non-relevant.

What are these relevant worlds then, so far, we've identified the following:

*Water*

$w_1$  The person next to me happily wins the lottery and leaves my drink toxin free.

$w_2$  The person next to me loses the lottery then spikes my drink with toxin.

*Flashes*

$w_3$  I was assigned to the non-experimental group, made to drink a pure orange juice, and was shown seven flashes.

$w_4$  I was assigned to the experimental group, was drugged and shown six flashes.

In  $w_1$  and  $w_3$ ,  $p$  is true. While in  $w_2$  and  $w_4$ ,  $p$  is false. The possible worlds considered as relevant in both cases are  $w_2$  and  $w_4$ . But are these worlds really closer to the actual one than  $w_1$  and  $w_3$ ? Obviously, the answer is no. Worlds  $w_1$  and  $w_3$  are more similar to the actual worlds described in both cases. Perhaps a little too similar, in fact my belief  $p$  is true in both worlds, like in the actual worlds. So, if we limit the set of close worlds to these worlds, my belief will be safe in both cases. However, if we limit the set of close worlds to worlds similar with respect to the truth of  $p$ , safety will be a trivial condition. On the other hand, if we include worlds 2 and 4 the counterexamples will hold.

*Atomic Clock (Bogardus)*

Tomas Bogardus offers this counterexample:

**ATOMIC CLOCK.** The world’s most accurate clock hangs in Smith’s office at a cereal factory, and Smith knows this. The clock’s accuracy is due to a clever radiation sensor, which keeps time by detecting the transition between two energy levels in cesium-133 atoms. This radiation sensor is very sensitive, however, and could easily malfunction if a radioactive isotope were to decay in the vicinity (a very unlikely event, given that Smith works in a cereal factory). This morning, against the odds, someone did in fact leave a small amount of a radioactive isotope near the world’s most accurate clock in Smith’s office. This alien isotope has a relatively short half-life, but—quite improbably—it has not yet decayed at all. It is 8:20 am. The alien isotope will decay at any moment, but it is indeterminate when exactly it will decay. Whenever it does, it will disrupt the clock’s sensor, and freeze the clock on the reading “8:22.” (Don’t ask why; it’s complicated.) Therefore, though it is currently functioning properly, the clock’s sensor is not safe. The clock is in danger of stopping at any moment, even while it currently continues to be the world’s most accurate clock. Smith is quite punctual, and virtually always arrives in her office on workdays between 8:20 and 8:25 am, though no particular time in that duration is more likely than any other to see her arrive. Upon entering her office, Smith always looks up at her clock and notes the time of her arrival. Today, in the actual world, that alien isotope has not yet decayed, and so the clock is running normally at 8:22 am when Smith enters her office. Smith takes a good hard look at the world’s most accurate clock—what she knows is an extremely well-designed clock that has never been tampered with—and forms the true belief that it is 8:22 am (*p*).<sup>35</sup>

---

<sup>35</sup> Bogardus, “Knowledge under Threat,” 12–13.



In *Atomic Clock*, Smith's belief  $p$  has several epistemic virtues. First, it is supported by evidence. Smith reasonably forms that belief after looking at a clock that is known to be the world's most accurate. Second, there is no *defeating evidence*. In fact, there is no defeater of any sort. Smith's belief is justified, true, and undefeated. Third, it's not grounded on any false belief. And lastly, it is formed via a reliable process. At 8:22 am in the actual world, the clock is still very accurate. These, among other things, warrant the intuition that Smith knows  $p$ .

But is Smith's belief safe? Bogardus says that it is not.<sup>36</sup> Remember that if the *alien isotope* decayed before or around the time Smith formed her belief, the clock would have malfunctioned, and her belief would have been false. And at the time Smith formed the belief the isotope is very likely to decay. So, there is a *close world* where the isotope decayed, the clock malfunctions and erroneously reads 8:22 am. In this possible world, Smith falsely believes that it's 8:22 ( $p$ ). Smith would have easily believed  $p$  even if it were false. So, Smith knows, but her belief is not safe.

Bogardus claims that *Atomic Clock* succeeds where other counterexamples failed, particularly those offered by Cosmesaña, Neta and Rohrbaugh, and Kelp (see my discussion of these cases above).<sup>37</sup> The difference is, in those examples, the subjects are no longer at epistemic risk when they formed their beliefs, while in *Atomic Clock*, the subject remains to be epistemically threatened at the time that she formed her belief.

Recall that in *Halloween* (Cosmesaña), *Water and Flashes* (Neta and Rohrbaugh), and *Clock* (Kelp), the subject *nearly* got into a situation where they would have falsely believed the proposition in question. But they actually avoided these situations. This happened in *Halloween* when Cosmesaña decided not to disguise himself as Michael; in *Water* when the person standing next to me won the lottery; in *Flashes*, when I was assigned to the control group; and in *Clock* when Russell came down the stairs at 8:22. In other words, in these examples, when the subject actually formed the belief in question, she was no longer in a situation where she could have falsely believed it. Bogardus's main contention is that: "One can be safe at  $t$  even if something nearly happened before  $t$  that would have put one in danger at  $t$ ."<sup>38</sup> He argues then that the beliefs in these examples are *safe*.

In contrast, the subject in *Atomic Clock* is in an actual situation where she could have easily formed a false belief. The threat is real and live, so to speak. At any time, the isotope could decay. It could have decayed before Smith came in, right before she looked at the clock, and even while she was looking at it. The clock could have easily malfunctioned. She could have

---

<sup>36</sup> *Ibid.*, 12.

<sup>37</sup> *Ibid.*, 16.

<sup>38</sup> *Ibid.*

easily falsely believed  $p$ . And she would have in a set of close possible worlds. And this makes her belief unsafe.

### *3/6 Clock (Freitag)*

Lastly, we turn to a case presented by Wolfgang Freitag:

**3/6 CLOCK.** The clock malfunctions . . . and shows either 3:00 or 6:00. It shows 3:00 at 3:00 and at all times between 6:00 and 11:58 (a.m. and p.m.). At all other times, it shows 6:00. Jim, not aware of the clock's malfunctioning, looks at the clock at 3:00, thereby picking up the true belief that it is 3:00 ( $p$ ).<sup>39</sup>

In *3/6 Clock*, we have another malfunctioning clock; it shows the right time only twice a day, at 3 in the morning and in the afternoon. The chance for Jim's belief to be true is only 1/360 given that the clock shows 3:00 only 12 hours a day. Intuitively then, Jim does not know  $p$ . But Freitag claims that Jim's belief is safe. Let me demonstrate his argument.

Note a few things first about *3/6 Clock*. First, what we have here is not your usual stopped clock. It shows 3:00 at 3:00 (a.m. and p.m.) and between 6:00 to 11:58 (a.m. and p.m.). From 3:01 to 5:59 (a.m. and p.m.) it shows 6:00, then again from 11:59 to 2:59 (a.m. and p.m.). Let's represent this on a table for easy reference:

TIME (A.M. and P.M.)	WHAT THE CLOCK SHOWS
11:59 – 2:59	6:00
3:00	3:00
3:01 – 5:59	6:00
6:00 – 11:58	3:00

So Jim, luckily, looks at the clock at 3:00 (whether it is a.m. or p.m. is not important since in either case the clock will show the correct time) and forms the true belief that *it is 3:00* ( $p$ ). Notice that Jim would have falsely believed  $p$  if he had looked at the clock at any time between 6:00 and 11:58. But, he would not have formed the same belief (and so not believe the same belief falsely) if he had looked at the clock a minute earlier or a minute later at 2:59 or at 3:01. In fact, the only other time he would have formed the same belief is between 6:00 and 11:58. If he looked at the clock at any other time, he would have formed an equally false but different belief, *it is 6:00* ( $q$ ).

<sup>39</sup> Freitag, "Safety, Sensitivity and 'Distant' Epistemic Luck," 11.

Now, from the qualifications given, we can identify at least four sets of possible worlds:

- w<sub>1</sub> 3:00 worlds
- w<sub>2</sub> 3:01–5:59 worlds
- w<sub>3</sub> 11:59–2:59 worlds
- w<sub>4</sub> 6:00–11:58 worlds

Recall again that Jim would have formed belief *p*, by looking at the clock, in w<sub>1</sub> and w<sub>4</sub>. Belief *p* is true in w<sub>1</sub>, and false in w<sub>4</sub> (call this the failure worlds). Jim would not have formed belief *p*, by looking at the clock, in w<sub>2</sub> and w<sub>3</sub>. Instead, he would have falsely believed *q* in these worlds. If we order these worlds in terms of similarity, *ceteris paribus*, the worlds close to the actual world clearly belongs to w<sub>2</sub> and w<sub>3</sub> (3:01 and 2:59 worlds). These worlds are so much closer to the actual world than any of the failure worlds, w<sub>4</sub> (6:00 and 11:58). Thus, Freitag claims that in the nearby worlds (w<sub>2</sub> and w<sub>3</sub>), Jim does not falsely believe *p*.<sup>40</sup> The subjunctive “*S would believe that p only if it were so that p*” obtains. Jim’s belief is safe. The failure worlds (worlds 4) are not close worlds.

Freitag shows that safety cannot account for cases that involve what he calls *distant* (non-close) *failure worlds*, provided that, *all things being equal*, in the nearest possible worlds, the subject would not have formed the same belief she forms in the actual world, using the same method she used in forming her belief in the actual world. Proponents of the modal theories can provide an analysis of *closeness* that would include the *distant* failure worlds in the set of close worlds. However, Freitag, similar to others (see discussion above), notes that it is difficult to provide a “consistent and convincing set of criteria” for closeness ranking.<sup>41</sup> And instead of tinkering with the given semantics or the intuitive similarity ordering, Freitag proposes that we fix safety by “searching for a different way of selecting relevant possible worlds.”<sup>42</sup>

## Conclusion

The counterexamples cited above further demonstrate why B-type and C-type beliefs fail the safety test. The dilemma as I already noted is that the conditional reliability or the instability of a method does not take away the epistemic worth of justified and true belief formed via an actually reliable

---

<sup>40</sup> *Ibid.*

<sup>41</sup> *Ibid.*, 16.

<sup>42</sup> *Ibid.*

method.<sup>43</sup> And since what makes the method actually reliable is a relevant epistemic detail, safety theories cannot respond to these objections by adjusting the similarity ranking without trivializing safety.

*Department of Philosophy, De La Salle University, Philippines*

## References

- Baumann, Peter, "Is Knowledge Safe?" in *American Philosophical Quarterly*, 45:1 (2008).
- Bogardus, Tomas, "Knowledge under Threat," in *Philosophy and Phenomenological Research*, 88:2 (2014).
- Cosmesaña, Juan, "Unsafe Knowledge," in *Synthese*, 146 (2005).
- Dacela, Mark Anthony L., "Are Modal Conditions Necessary for Knowledge?" in *Kritike: An Online Journal of Philosophy* 13:1 (2019).
- Freitag, Wolfgang, "Safety, Sensitivity and 'Distant' Epistemic Luck," in *Theoria*, 80:1 (2014).
- Goldman, Alvin I., "What is Justified Belief?" in *Justification and Knowledge*, Philosophical Studies Series in Philosophy, Vol. 17, ed. by George Sotiros Pappas (Dordrecht: Springer, 1979), 1–23.
- Kelp, Christoph, "Knowledge and Safety," in *Journal of Philosophical Research* 34 (2009).
- Lewis, David, "Counterfactual Dependence and Time's Arrow," in *Noûs* 13:4 (1979).
- \_\_\_\_\_, *Counterfactuals* (Oxford: Blackwell Publishers, 1973).
- Neta, Ram and Guy Rohrbaugh, "Luminosity and the Safety of Knowledge," in *Pacific Philosophical Quarterly* 85 (2004).
- Sosa, Ernest, "How to defeat Opposition to Moore," in *Philosophical Perspectives* 13 (1999).
- Stalnaker, Robert, "A Theory of Conditionals," in *Studies in Logical Theory*, ed. by Nicholas Rescher (Oxford: Basil Blackwell, 1968), 98–112.
- Steup, Matthias, *An Introduction to Contemporary Epistemology*. Upper Saddle River, NJ: Prentice-Hall, 1996).

---

<sup>43</sup> Dacela, "Are Modal Conditions Necessary for Knowledge?" 114.

## **Towards an Experimental Turn in Filipino Philosophy: A New Way Forward**

---

*Ian Anthony B. Davatos*

**Abstract:** The primary objective of this paper is to find out whether there is any possibility of coming up with a philosophy that we can call Filipino. Inspired by the works of Leonardo Mercado, I suggest an exciting new area of philosophy that can get us to an answer: experimental philosophy. Secondly, I shall bridge the connection between experimental philosophy and the search for Filipino philosophy. More specifically, I shall provide an answer as to how experimental philosophy can be expected to lead to a Filipino philosophy. Then, I shall suggest a novel way in how to do experimental Filipino philosophy, that is, experimental philosophy in the service of discovering a Filipino philosophy, and it is by way of traditional empirical methods in anthropology, such as interviews and focus group discussions. Finally, I introduce the charge of limited applicability inspired by Roland Theuas Pada and respond to the objection. I conclude by inviting Filipino philosophers to integrate experimental philosophy in their search for a Filipino philosophy.

**Keywords:** Mercado, Pada, Filipino philosophy, experimental philosophy

### **Introduction**

**D**o we need another discussion of Filipino philosophy? It seems that the time for debates has long been exhausted with no sign of being settled. While there are a few pioneers in the search for a genuine Filipino philosophy, many current Filipino philosophers seem content with just doing philosophy without the designation of “Filipino” before it. When asked whether he was trying to build a Filipino philosophy, Roque Ferriols, a prominent Filipino philosopher and metaphysician, answered a resolute no. He said:

No one can create a Filipino or anything else philosophy except by accident. Chuang Tzu did not try to develop a Chinese philosophy. He simply awoke to the Way within him and around him, tried to awake even more, knew that what he lived could not be put into words—when all that can be said has been said, the most important thing cannot be said—yet felt compelled to say all that he could say. Hundreds of years later what he said still lives and is called Chinese philosophy. He is surprised. It is the Way that matters to him, not the label.<sup>1</sup>

That there cannot be a Filipino philosophy (except by accident) is an assumption that I suspect continues to pervade the minds of many Filipino philosophers. This assumption is not totally unwarranted. For one, the label of Filipino philosophy has a relativistic ring to it. If one is attempting to construct a grand metaphysics, it seems ill-advised to call it Filipino metaphysics as if it is a view of reality that can only apply to Filipinos. If one is trying to advance a view of reality that is rationally convincing, one must make sure that it is a view that should appeal to people from all walks of life, Filipinos or otherwise. Any discovery of a Filipino philosophy seems too narrow in scope and too restrictive to be universally appealing.

Thus, despite some attempts to uncover a Filipino philosophy, philosophy in the Philippines is largely done in a way that follows the typical standard, which is generally a Western one. While many Filipino philosophers may insist that they are simply doing philosophy, they cannot deny that much of the assumptions, questions, and arguments with which they interact in their thinking and works bear the influence of Western thinking. Of course, it should not be denied that Western philosophy has gifted us with concepts, principles, and arguments that can be expected to command universal assent insofar as one possesses reason.<sup>2</sup> However, should it be correctly assumed that it is mostly the West which can claim the authority to command what should count as philosophy? By contrast, is it not possible by design, not by accident, as Ferriols assumes, to uncover a truly genuine Filipino philosophy, regardless of whether it is universally

---

<sup>1</sup> Roque J. Ferriols, "A Memoir of Six Years," in *Philippine Studies*, 22 (1974), 339.

<sup>2</sup> An anonymous reviewer pointed out that the whole premise of philosophy is already Western to begin with and suggested different naming conventions in our intellectual discourse, much like what is done by the Indians and the Chinese. This suggestion is exactly what will make the use of experimental philosophy relevant and philosophically interesting (as will be argued shortly): the discovery of a new set of intuitions from the Filipino people may initiate novel naming conventions that are not totally held captive by Western ideas.

persuasive or not? In this paper, I plan to focus on the second question by attempting to answer it in the affirmative. It is inevitable, however, that whatever one answers in the second question will have ramifications in how one answers the first.

As I shall argue, a fruitful first step in this project is to make use of experimental philosophy. Sometimes called *x-phi*, experimental philosophy is a new movement in contemporary analytic philosophy that makes use of empirical methods, especially experimental methods in psychology, in order to illuminate philosophical questions. As Eugen Fischer and John Collins explain, “Experimental philosophers use empirical surveys and experiments to develop an understanding of philosophically relevant intuitions that helps us determine whether we should accept or reject them.”<sup>3</sup> Two things are to be noted from this.

First is the element of empirical surveys and experiments. While many varying accounts have been advanced to delineate the meaning of experimental philosophy,<sup>4</sup> experimental philosophers are united in the use of empirical methods in doing experimental philosophy. The methods commonly utilized are those of psychology, especially controlled experiments; although some have counted philosophically motivated ethnography as a method of experimental philosophy.<sup>5</sup> This form of philosophy is sure to diverge from the common way of doing philosophy that is done from the armchair, that is, philosophy that makes use of *a priori* principles and intuitions in order to argue for or against a philosophical position. It is called armchair philosophy as such because it is done in the comfort of one’s armchair as it were with little need for a fieldwork or empirical grounding to one’s argumentation. This particular description of philosophy is not meant to be disparaging but is simply a statement of a dominant practice within the discipline. A well-known example of armchair philosophizing is the so-called analytic tradition in philosophy, especially as it is currently practiced in the Western philosophical arena.<sup>6</sup> This tradition is known for making use of conceptual analysis, where a certain concept is

---

<sup>3</sup> Eugen Fischer and John Collins, “Introduction,” in *Experimental Philosophy, Rationalism, and Naturalism: Rethinking Philosophical Method*, ed. by Eugen Fischer and John Collins (New York: Routledge, 2015), 4.

<sup>4</sup> Joshua Knobe and Shaun Nichols, “An Experimental Philosophy Manifesto,” in *Experimental Philosophy*, Vol. 1, ed. by Joshua Knobe and Shaun Nichols (New York: Oxford University Press, 2008), 3–14; David Rose and David Anks, “In Defense of a Broad Conception of Experimental Philosophy,” in *Metaphilosophy*, 44 (2013), 512–532; Joshua Alexander, *Experimental Philosophy: An Introduction* (Cambridge, UK: Polity Press, 2012).

<sup>5</sup> Stephen Stich and Kevin P. Tobia, “Experimental Philosophy and the Philosophical Tradition,” in *A Companion to Experimental Philosophy*, ed. by Justin Sytsma and Wesley Buckwalter (Oxford: Blackwell, 2016), 5.

<sup>6</sup> I thank one of the anonymous reviewers for asking me to provide this.



analyzed by identifying its necessary and sufficient conditions, and the conditions identified will then be applied to hypothetical scenarios to see whether they satisfy what is supposed to be the true meaning of the concept. Also, this tradition is known for making use of contemporary technical tools, such as symbolic logic and probability theory, especially as it is applied in rigorous argumentation.

Second important element is the role of intuition in philosophizing, and how experimental philosophy serves to examine the extent to which philosophical intuitions track truth. Before experimental philosophy, it is commonly assumed that intuitions in philosophy are good indicators of philosophical truths or at least a particular theory's overall (im)plausibility, especially if those of the philosopher's agree with those of the layman's. Stephen Stich and Kevin Tobia describes this process well:

A philosopher describes a situation, sometimes real but more often imaginary, and asks whether some of the people or objects or events in the situation described have some philosophically interesting property or relation .... When things go well, both the philosopher and her audience will agree on an answer, with little or no conscious reflection, and they will take the answer to be *obvious*. The answer will then be used as evidence for or against some philosophical thesis. The mental states that underlie episodes of this sort are paradigm cases of philosophical intuitions.<sup>7</sup>

The imaginary situation that is mentioned above refers to so-called thought-experiments in philosophy, which seeks to elicit the desired intuitions from others in order to argue for or against a particular philosophical view. This has been known as the "Method of Cases."<sup>8</sup> The role of intuition is highlighted when one looks at the frequency by which thought experiments are used throughout the history of philosophy and among the various branches of philosophy.<sup>9</sup> Stich and Tobia, for instance, have collected

---

<sup>7</sup> *Ibid.*, 6.

<sup>8</sup> Elizabeth O'Neill and Edouard Machery, "Experimental Philosophy: What is it Good For?" in *Current Controversies in Experimental Philosophy*, ed. by Edouard Machery (New York: Routledge, 2014), xiii; Max Deutsch, "Gettier's Method," in *Advances in Experimental Philosophy and Philosophical Methodology*, ed. by Jennifer Nado (New York: Bloomsbury, 2016), 69; Edouard Machery, "The Illusion of Expertise," in *Experimental Philosophy, Rationalism, and Naturalism: Rethinking Philosophical Method*, ed. by Eugen Fischer and John Collins (New York: Routledge, 2015), 189–193.

<sup>9</sup> Michael T. Stuart, Yiftach Fehige, and James Robert Brown, *The Routledge Companion to Thought Experiments* (New York: Routledge, 2018).

an impressive, albeit brief, catalog of thought experiments in contemporary analytic philosophy.<sup>10</sup>

In order to delve deeply into how essential intuitions are in the project of experimental philosophy, it is necessary that we look into the two programs of experimental philosophy. The so-called *negative program* aims to explore how reliable intuitions are in tracking truth. For most of contemporary analytic philosophy, the assumption has been that the contents of philosophical intuitions are probably true. Connected to this is the assumption that the intuitions of professional philosophers are universally shared across different cultures and demographics. However, there are empirical evidences that seem to undermine both these assumptions. Experimental studies have shown that many of the commonly held intuitions that underwrite certain philosophical views are subject to influences that are irrelevant to their truth. A critic of experimental philosophy, Max Deutsch recognizes the alleged success of the negative program in discovering what he calls the “truth-irrelevant variability in philosophical intuitions.”<sup>11</sup> Related to that, intuitions have been found to vary when one factors in gender,<sup>12</sup> personality,<sup>13</sup> and language.<sup>14</sup> With the advent of these empirical discoveries, it is no longer obvious to claim that philosophical intuitions appealed to by philosophers are universally shared and are likely to be true. And if these empirical studies are indeed successful in showing that intuitions are based on irrelevant factors, experimental philosophy in the guise of its negative program poses a great challenge to the traditional way of doing philosophy that makes use of intuitions in determining the truth about philosophical issues and concepts.<sup>15</sup>

One might get the idea that x-phi aims to undermine the use of intuition *per se*, but this is a misunderstanding. Experimental philosophers do

---

<sup>10</sup> Stich and Tobia, “Experimental Philosophy and the Philosophical Tradition,” 7.

<sup>11</sup> Max Deutsch, *The Myth of the Intuitive: Experimental Philosophy and Philosophical Method* (Cambridge: The MIT Press, 2015), 17.

<sup>12</sup> Wesley Buckwalter and Stephen Stich, “Gender and Philosophical Intuition,” in *Experimental Philosophy*, Vol. 2, ed. by Joshua Knobe and Shaun Nichols (New York: Oxford University Press, 2013).

<sup>13</sup> Adam Feltz and Edward T. Cokely, “Do Judgments about Freedom and Responsibility Depend on Who You Are? Personality Differences in Intuitions about Compatibilism and Incompatibilism,” in *Consciousness and Cognition*, 18:1 (2009), 342–350.

<sup>14</sup> Albert Costa, Alice Foucart, Sayuri Hayakawa, Melina Aparici, Jose Apestequia, Joy Heafner, and Boaz Keysar, “Your Morals Depend on Language,” in *PLoS ONE*, 9:4 (2014), e94842; Edouard Machery, Christopher Y. Olivola, and Molly De Blanc, “Linguistic and Metalinguistic Intuitions in the Philosophy of Language,” in *Analysis*, 69 (2009), 689–694.

<sup>15</sup> There have been serious doubts, however, on how successful these empirical discoveries are in showing what they purport to show. See for example Kaija Mortensen and Jennifer Nagel, “Armchair-Friendly Experimental Philosophy,” in *A Companion to Experimental Philosophy*, ed. by Justin Sytsma and Wesley Buckwalter (Oxford: Blackwell, 2016), 58–60.

not intend on discarding intuition for good and replace it solely with controlled experiments and empirical surveys. Rather, one of their major aims in line with the negative program is the identification of philosophically insignificant differences that influence what intuitions will be appealed at on any given time. As experimental philosopher Jonathan Weinberg clarifies, it is a misconstrual to view the negative program of x-phi as an assault on intuitions, full stop. Rather, “the target of the negative program has always been an armchair-based intuitive methodology and not intuitions *tout court*.”<sup>16</sup>

What is commonly assumed in the discussion is that such factors as personality, gender, and language seem to endanger the truth-tracking capacity that philosophers have long attributed to intuition since these factors seem irrelevant to the truth of a philosophical view. In their manifesto, pioneering experimental philosophers Joshua Knobe and Shaun Nichols have this to say about such discovery: “If I find out that my philosophical intuitions are a product of my cultural upbringing, then, since it’s in some sense an accident that I had the cultural upbringing that I did, I am forced to wonder whether my intuitions are superior at tracking the nature of the world, the mind, and the good.”<sup>17</sup> The same goes with gender or any other external factors: If I discover that my intuitive attractions to a particular moral theory are largely due to its “masculine” elements (and I am a male), then I have reason to suspend my belief that the theory is true.

While the negative program of x-phi is clear in its critical assessments of philosophical intuitions, especially in their supposed role of supporting philosophical views, x-phi also has its *positive program*. This side of x-phi aims to explore intuitions experimentally in order to improve conceptual analysis, which is a major defining project for contemporary analytic philosophy. A major goal of the positive program is to avoid philosophical echo-chambers in which professional philosophers rely on their and their colleagues’ intuitions that may have been rooted in theoretical commitments absent in the intuitions of non-philosophers. Thus, the folk intuitions of non-philosophers are taken into account in order to generate discoveries that may be relevant in illuminating philosophical issues and problems. An example is in order.<sup>18</sup> In a recent study, an interdisciplinary team composed of two psychologists, a professional counselor, a philosopher, and a theologian investigated a number of Christians who have had cancer experience, and how they interpreted such experience as it relates to their belief in God. While

---

<sup>16</sup> Jonathan M. Weinberg, “Going Positive by Going Negative: On Keeping X-Phi Relevant and Dangerous,” in *A Companion to Experimental Philosophy*, ed. by Justin Sytsma and Wesley Buckwalter (Oxford: Blackwell, 2016), 72.

<sup>17</sup> Knobe and Nichols, “An Experimental Philosophy Manifesto,” 11.

<sup>18</sup> I thank one of the anonymous reviewers for asking me to provide this.

there are a good number of interesting points from the study, one intriguing result came out, which is that the research participants intuit a position towards their experience of evil that is not well-known in the literature on theodicy, and it is the idea of trusting God in mystery.<sup>19</sup> This idea is a ripe topic for further defense and study, and a discovery such as this is a fruit of the positive program of experimental philosophy.

With these two programs in mind, let us now turn to the project and intent of Leonardo Mercado in discovering Filipino philosophy and how such a project can be assisted by experimental philosophy as a worthwhile intellectual endeavor.

### Experimental Philosophy in the Service of Filipino Philosophy

When Filipino philosopher Leonardo Mercado first came into the scene with his major work titled *Elements of Filipino Philosophy*, he recalled that it was “met with skepticism.”<sup>20</sup> This doubt has been re-echoed by many scholars who are familiar with Mercado’s work. While Emerita Quito commended Mercado for his pioneering attempt to establish a Filipino philosophy, she admitted that there was much opposition to his theories. As she said, “Scholars find his work to be merely linguistic.”<sup>21</sup>

How did Mercado approach his works that command this judgment? This is because Mercado was the first to form a pathway that made use of an approach—particularly of linguistics—that is not purely philosophical to unearth a Filipino philosophy. Since his method was focused generally on analyzing the intricacies of certain languages, it is hasty to claim that Mercado’s works utilized strictly empirical methods, ones that are used by professional anthropologists and psychologists when doing their fieldwork. But the point of Mercado’s initiative had been to use certain methodologies in science in order to discover a systematic form of Filipino philosophy. This atypical approach in philosophizing was expected to draw some negative impressions from Filipino philosophers who have learned to practice philosophy from the armchair, that is, discovering *a priori* principles and intuitions, and thereby using such intuitions to generate philosophical

---

<sup>19</sup> Jason Silverman, Elizabeth Hall, Jamie Aten, Laura Shannonhouse, and Jason McMartin, “Christian Lay Theodicy and the Cancer Experience,” in *Journal of Analytic Theology*, 8 (2020), 359–61.

<sup>20</sup> Leonardo Mercado, “Reflections on the Status of Filipino Philosophy,” in *Kritike: An Online Journal of Philosophy*, 10:2 (2016), 21.

<sup>21</sup> Emerita S. Quito, “The Filipino and the Japanese Experience,” in *Lectures on Comparative Philosophy* (Manila: De la Salle University, 1979), 34, as quoted in Emmanuel D. Batoon, “Tracing Mercado’s Anthropological Perspective (Second of Two Parts),” in *Kritike: An Online Journal of Philosophy*, 8:2 (December 2014), 2.

insights. This is known as the deductive approach, which according to Mercado, is exemplified by the Western mind.<sup>22</sup>

Unlike the West, the Filipino mind thinks inductively and intuitively, a claim which Mercado tried to prove by analyzing the Filipino language through poetry, proverbs, and the *balagtasan*.<sup>23</sup> The use of proverbs and *balagtasan* as a form of debate are two ways that show that Filipinos think more inductively rather than deductively. For Mercado, Filipinos are poetic in their reasoning. According to Mercado, there is nothing wrong with this because poetry and the emphasis on the concrete can similarly reach the truth as much as abstract logic can. As Mercado said, “poetic symbols can serve as paradigms for intuition.”<sup>24</sup> The important thing to note here is the reliance on intuition as it is discovered in language and the belief that intuition categorically establishes what counts as basic and fundamental truths. He called his basic approach the metalinguistic approach, which “rests on the supposition that a language mirrors the thought and somehow determines the outlook of its native speakers.”<sup>25</sup>

Mercado’s attempt to discover a Filipino philosophy had been tied to his goal of separating Filipino thought from the pervasive influence of Western thinking. His worry had been that for as long as Filipino thinkers are tied to the paradigm of Western thinking, the difficulty of having our own Filipino philosophy will always fall by the wayside. One might retort that Mercado’s methodology is also derived from Western theories, which makes his claim seem ironic.<sup>26</sup> This is true, and this objection underlies the fact that we cannot completely expunge our way of thinking from some form of Western influence. But this admittance does not entail that there cannot be forms of Filipino thinking that are unique and separable from Western influence, especially since what was appropriated by Mercado had been the methodology used, something that may be separated from the most basic intuitions of (Filipino) philosophical thinking.

But why should there be a need to discover a Filipino philosophy, if there is such a thing at all? Is it not enough to philosophize, as Ferriols suggested, and to let history decide whether the fruits of one’s philosophical labor are worthy of the name “Filipino philosophy”? Bear in mind that Mercado viewed philosophy not as “a huge shell game, a Brainiac sport played hard just for the fun and posturing of it,” as Damien Broderick would

---

<sup>22</sup> Leonardo Mercado, *Elements of Filipino Philosophy* (Tacloban: Divine Word University Publications, 1974), 73–91.

<sup>23</sup> Leonardo Mercado, “Reasoning,” in *The Filipino Mind* (Washington, D.C.: The Council for Research in Values and Philosophy, 1994), 41–54.

<sup>24</sup> *Ibid.*, 51.

<sup>25</sup> Leonardo Mercado, “Filipino Thought,” in *Philippine Studies*, 20:2 (1972), 207.

<sup>26</sup> I thank an anonymous reviewer for raising this objection.

colorfully describe it.<sup>27</sup> As Mercado said, “The love of wisdom is not just a leisurely search for truth.”<sup>28</sup> Philosophy should be socially relevant, and Mercado viewed this relevance in terms of how philosophy can aid in erecting a strong foundation for Filipino identity, an identity that should be made separate from the contours of Western intellectual influence. By assisting to build a robust Filipino identity, philosophical thinking becomes a catalyst to nation-building and a healthy nationalism. In an interview with Emmanuel C. De Leon and Marvin Einstein S. Mejaro, Mercado was clear that his philosophy is “a form of nationalism.”<sup>29</sup> Many Filipino philosophers are resistant to his way of doing philosophy because according to Mercado many Filipino philosophers of today are still enamored by Western philosophical legacies. According to Mercado, the colonial mentality exhibited by this tendency will not do good for the country.<sup>30</sup>

So, it should be clear by now why there is a need to discover a Filipino philosophy: it is because having a philosophy that we can consider Filipino is an essential tool towards intellectual nationalism. It is not only important to be a Filipino citizen whose country is now independent from colonizers, but to develop and appreciate the way(s) by which the Filipino *as Filipino* thinks, especially about perennial philosophical issues. While many are still skeptical that Mercado indeed uncovered elements of Filipino philosophy, I think that the works of Mercado have borne fruit in terms of revealing how certain elements in Filipino thinking may diverge from the paradigm of Western thinking. I have already mentioned how Mercado differentiated between Filipino and Western forms of reasoning by looking at some examples of Filipino poetry, proverbs, and debate. Mercado also examined Filipinos’ views about soul, beauty, and evil, among others—views that are embedded in different Philippine languages.<sup>31</sup>

What is more interesting about Mercado’s works is how they suggest that there may be Filipino intuitions that are distinct from those of the West. This is an exciting discovery and precisely because this may be strengthened by some of the pioneering discoveries of experimental philosophy. In the early days of experimental philosophy, one of the most staggering results is

---

<sup>27</sup> Damien Broderick, “Introduction II: Philosophy on the Inclined Plane,” in *Philosophy’s Future: The Problem of Philosophical Progress*, ed. by Russell Blackford and Damien Broderick (New York: Wiley-Blackwell, 2017), 13.

<sup>28</sup> Leonardo Mercado, “What is Philosophy?” in *Filipino Thought* (Manila: Logos Publications, Inc., 2000), 9.

<sup>29</sup> Emmanuel De Leon and Marvin Einstein Mejaro, “An Interview with Leonardo Nieva Mercado, SVD,” in *Kritike: An Online Journal of Philosophy*, 10:2 (December 2016), 4.

<sup>30</sup> Mercado, “What is Philosophy?” 9.

<sup>31</sup> Leonardo Mercado, *The Filipino Mind* (Washington D.C., The Council for Research in Values and Philosophy, 1994).



how intuitions differ from culture to culture.<sup>32</sup> What may be taken as an intuitive truth of Westerners may not be the case for East Asians, and this is similar to what Mercado's works implied since he first published his *Elements: Filipino philosophy*, or at least some of its elements, grounded in the intuitions of Filipinos may be distinct from Western philosophy insofar as the latter is grounded in intuitions of the West. What is more exciting is the fact that in the experimental studies regarding culture, what are referred to as East Asians are Chinese and Japanese and never Filipino. For instance, in a book about "how Asians and Westerners think differently and why," the author Richard E. Nisbett did not mention Filipinos even once.<sup>33</sup> This gap is also reflected in other experimental studies.<sup>34</sup> Once the implication of this lacuna is adequately recognized, it is astonishing how much vast the research terrain is available for Filipino philosophers to analyze and study. Thus, experimental philosophy can clarify to philosophers in search of Filipino philosophy some methodological pathways by which they can proceed and are likely to succeed. Indeed, Mercado has just opened up the floodgate into a rich, wide-ranging and underexplored territory that is Filipino philosophy.

Now, if there is indeed some promise with using experimental philosophy on our search for a Filipino philosophy as I have so far argued, then it is time that we proceed to laying out the details on how exactly Filipino philosophy is to be discovered through  $\chi$ -phi. By the term itself, experimental philosophy is a method of philosophizing that makes use of controlled experiments and other quantitative methods in illuminating philosophical issues. There is nothing wrong with this but using this quantitative-experimental approach as the starting point may preempt what and how philosophical issues are considered by the participants, which in our case would be Filipinos, especially the non-philosophers. To address this concern, it is sometimes more appropriate to start with qualitative research methods, such as in-depth interviews and focus group discussions.

The point of qualitative research methods is to avoid for as much as possible the tendency of the interviewer to impose their preconceived notions

---

<sup>32</sup> Jonathan Weinberg, Shaun Nichols, and Stephen Stich, "Normativity and Epistemic Intuitions," in *Philosophical Topics*, 29:1/2 (2001), 429–460; Edouard Machery, Ron Mallon, Shaun Nichols, and Stephen Stich, "Semantics, Cross-Cultural Style," in *Cognition*, 92:3 (2004), B1–B12; Linda Abarbanell and Marc D. Hauser, "Mayan Morality: An Exploration of Permissible Harms," *Cognition*, 115 (2010), 207–224; Henrick Ahlenius and Torbjörn Tännsjö, "Chinese and Westerners Respond Differently to the Trolley Dilemmas," in *Journal of Cognition and Culture*, 12:3–4 (2012), 195–201.

<sup>33</sup> Richard E. Nisbett, *The Geography of Thought: How Asians and Westerners Think Differently ... and Why* (New York: The Free Press, 2003).

<sup>34</sup> Joseph Henrich, Steven J. Heine, and Ara Norenzayan, "The Weirdest People in the World?" in *Behavioral and Brain Sciences*, 33:2–3 (2010), 61–83; Machery, Mallon, Nichols, and Stich, "Semantics, Cross-Cultural Style," B1–B12; Hazel Markus and Shinobu Kitayama, "Culture and the Self," in *Psychological Review*, 98 (1991), 224–253.



on the topic under discussion. This is in direct contrast with quantitative research that makes use of statistics since the latter is utilized to confirm or refute a particular hypothesis. Of course, both methods are important but by starting our project with the use of qualitative methods, the imposition of our theoretical commitments is minimized and the discovery of novel ways of thinking from the participants themselves is encouraged.

How should this be actually done? We start with a philosophically rich concept that is already understandable to the participant. We need to bear in mind that the point of the interview is to explore the various ways by which the participant, *not the interviewer*, understands the concept. So what is suggested is a highly unstructured interview; it is structured only insofar as it starts with a certain philosophical concept but the trajectory of the interview should be as free-flowing as possible. The interviewer is there only to help the participant explore the concept according to the latter's understanding. It is important for the interviewer to avoid imposing her theoretical understanding on the matter and lead the interview to the direction to which she wants it to go. This advice of caution is essential if one wants to get the most out of the interview and by that I mean that there may be some insights that the participant has with regard a philosophical idea which may have been overlooked due to the ways in which the community of professional philosophers initially approached the topic. By letting the participant speak for himself, with little philosophical influence from the interviewer, the participant may come to express his view about the topic in directions that the interviewer may not expect. This is neither to say that philosophy is fully expunged from the discourse<sup>35</sup> nor to claim that the participant does not have any philosophical influences. It is simply to allow the possibility that the participant may have some philosophically interesting ideas that have not been entertained or fully elucidated by philosophers. The important thing to note here is that a qualitative interview involves open response answers. Unlike quantitative interviews that usually require a yes or no answer or answer that fits within a set of choices, questions in qualitative interviews "give participants the opportunity to provide an open response in their own terms."<sup>36</sup>

Of course, this approach to the interview should not be viewed as an easy task. The interviewer, who is assumed as a philosopher by profession, has a set of related concepts, arguments and schools of thought in her intellectual toolkit. This arsenal of specialized knowledge is a natural strength of philosophers but this may impede certain interesting discoveries from an interview by way of leading the interview to issues that are unfamiliar to the

---

<sup>35</sup> I thank an anonymous reviewer for raising this concern.

<sup>36</sup> James Andow, "Qualitative Tools and Experimental Philosophy," in *Philosophical Psychology*, 29:8 (2016), 1129.

participant. To address this, professional philosophers need careful advice, professional guidance, and experience in the field. The first two may be acquired through collaboration with other disciplines such as anthropology, psychology, and other social sciences, a step which is now a staple of many works in experimental philosophy. Experience in the field requires time, effort, and peer assessment in order to be fruitful, but a considerable time in the field is necessary in order to master the forms of interview that will best achieve the interviewer's goals. Since our focus here is interview, one may especially learn from anthropologists and psychologists who are experts in qualitative methodologies such as qualitative interviewing.

Aside from the interview, there is also a focus group discussion (FGD) as another common qualitative method of research. In the interview, the major goal is to explore the depths by which the participant understands and makes sense of the concept. The interviewer is present only to motivate the participant in exploring further the latter's general understanding of a concept; it should not be a venue for the interviewer to try to confirm whatever philosophical position she may have. So generally, the aim of the interview is to uncover reflective thinking of the participants. It is meant to identify the various reasoning processes involved in thinking about a philosophical topic. However, in an FGD, a participant is no longer alone with the interviewer but in the companion of people who may or may not agree with his point of view. This methodology, therefore, may be used to probe further a participant's reasoning processes. But again, one may also begin with an FGD with the experimental philosopher as the facilitator and then using the findings in the said group discussion as starting points when one eventually decides to proceed to the qualitative interview. A participant may have said something philosophically interesting in the FGDs that is not explored in further detail due to the nature of FGDs (where a number of participants are involved and majority likes to participate and speak). This is where the follow-up individual interview shows its strength. In any case, whether one begins with an FGD then an interview or vice versa, either way is methodologically viable.

Qualitative methodologies, such as interview and FGD, have transcripts as its main data. The average time for an individual interview is around 30 minutes to an hour depending on the competence of the interviewer; while it is around 2 to 3 hours for FGDs. Once transcribed, 10 interviews and even just one FGD can produce hundreds upon hundreds of transcript pages. Analysis of such a huge data requires technical competence that is expectedly unfamiliar to a philosopher with no experience in social science research. James Andow, the first philosopher to argue for the use of qualitative methodologies in experimental philosophy, has a lot to say about

the process of qualitative data analysis, and I refer the reader to his pioneering article for more details on the qualitative process.<sup>37</sup>

The major point of using a qualitative approach in experimental philosophy is to uncover the reasoning process by which the participants make sense of the concept in question. What is important to discover is the conceptual structure of the participants involved in understanding a philosophical term. For instance, it is fascinating to hear many Filipino parents express that at some point they will no longer micromanage their children's behaviors since "they already know what's right and wrong." From the moral-psychological perspective, this assumption is intriguing. How do the parents know that their children are at that point in their lives where they already know the contents of morality? In what contexts does this utterance arise? A qualitative methodology can probe deeper into this familiar claim and by doing so, experimental philosophers may discover new insights into moral epistemology especially as it viewed by Filipino parents and even by their children. There may be a wellspring of related, even novel, moral concepts, arguments, and territories waiting to be explored within the vicinity of Filipino moral epistemology and Filipino moral psychology among many others. However, as intimated before, the interviewer should begin with a familiar concept, such as morality, and find way to incorporate such notion as moral epistemology (and a host of others) within the broader notion of morality in the course of the interview or focus group discussion. This defense of qualitative methods in x-phi is in line with Andow's view. According to him, the point of qualitative methods in experimental philosophy is to discover "the reflective aspects of ordinary thought about philosophically interesting things" since philosophers draw on them in their works.<sup>38</sup> Andow then lays out several ways by which the reflective thinking of the ordinary folk may be philosophically valuable, such as enriching or challenging the philosopher's evidence base.<sup>39</sup>

Aside from unearthing the conceptual structure embedded in the participant's understanding of a term, qualitative methodologies may also be used to uncover the most fundamental intuitions that Filipinos have regarding philosophical concepts in question. This diverges from James Andow's defense of qualitative methods in experimental philosophy. For him, there is good reason to think that "qualitative methods are unsuitable for measuring the ordinary, subpersonal mechanisms underlying processes like moral judgment."<sup>40</sup> However, I do not agree with James Andow in his claim that if x-phi aims to study intuitions, then qualitative methods have

---

<sup>37</sup> Andow, "Qualitative Tools and Experimental Philosophy."

<sup>38</sup> *Ibid.*, 1134.

<sup>39</sup> *Ibid.*, 1135–1136.

<sup>40</sup> *Ibid.*, 1133.

little to contribute to it. While he is right in his view that qualitative methods seek to uncover reflective thinking, the role of intuitions in this form of thinking cannot be easily disregarded since they may be utilized to discover what intuitions serve as the primary backbone in the rationalization of one's position. How exactly can this be done?

If we will follow the common characterization of intuition suggested by Andow as "non-inferential judgements that are not a product of conscious reasoning, are fairly immediate, and not slowly or carefully reasoned,"<sup>41</sup> then intuitions may emerge in an interview when the participants reach certain propositions of their reasoning process that they take to be simply properly basic. As such, intuitions here are construed as the basic blocks of a reasoning process which the participants take to be so obvious as not needing of further defense. Recognizing the point where the participants manifests their basic intuitions in the interview or FGD is a skill that the researcher needs to develop and be constantly mindful of. In fact, it would enhance the authenticity of the data if the researcher would note and include in her analysis detailed observations of the participants as the latter tries to justify what appears to them as commonsensical or even universally accepted proposition.

On that note, it need not always be the case that we start with a qualitative, rather than a quantitative approach in experimental philosophy. While quantitative methods in the form of surveys are largely used to uncover intuitions and the psychological processes that underlie those intuitions, they may also be utilized to identify certain approaches to a philosophical topic or patterns of thinking underlying a position in a philosophical issue that call for deeper probing. Aside from just uncovering philosophical intuitions (as how experimental philosophers envision quantitative methods to be), they may suggest topics for further philosophical study that are initially elucidated using a qualitative approach.

### **Filipino Philosophy and the Charge of Limited Applicability**

If the argument just laid out is successful, it provides an unambiguous way of discovering Filipino philosophy, which is by making use of the tools of experimental philosophy. However, by narrowing the contents of Filipino philosophy to the philosophical intuitions of Filipinos, this might invite the charge that Filipino philosophy as so far conceived here, is of limited applicability. After all, if Filipino philosophy is composed generally of ideas and arguments whose foundations are the philosophical intuitions of Filipinos—not American, German, or British—then, one may

---

<sup>41</sup> *Ibid.*

easily conclude that such arguments will only appeal and persuade Filipinos. One of the representatives for this charge is Roland Theuas Pada in his not-so-recent article in this journal.<sup>42</sup>

At the outset, Pada is to be commended for giving justice to the projects of Filipino philosophers, such as Leonardo Mercado and Florentino Timbreza, even if he generally disagrees with them. He has described their projects as “a quasi-ethnological attempt to construct a unified description of Filipino thought through various ethnic practices.”<sup>43</sup> However, what Pada wants to argue is that a variant of Filipino philosophy that is restricted to the ethnological descriptions of culture—such as that of Mercado and Timbreza—will fall short of the practice of philosophy as a discourse. As he remarks, “Isolating our idea of what a ‘grassroot’ philosophy is as an ethnic practice fails the implicit criteria of philosophy as a constantly continuing discourse, which I think is not only detrimental to the idea of what philosophy is, but is also against the idea of *philosophy* as a *discourse*.”<sup>44</sup> For Pada, when the criterion of discourse is applied in the practice of philosophy, the result is a new category: “the development of a discursive philosophy that originates from Filipino thinkers and engages with the tradition of philosophy as a whole.”<sup>45</sup> Pada sees then the projects that are “strictly limited to the national or cultural concerns of their own life-world” as needing expansion to produce “works that are read, not because of their national origin, but because of their effect to philosophy in general.”<sup>46</sup>

Central to Pada’s view of philosophy is the idea of discourse. He does not deny that the approach of Mercado and Timbreza can generate philosophical interest in the local scene, which for him is helpful if we want to develop a strong grassroots tradition.<sup>47</sup> With this, he is not far from the nationalistic goal that Mercado envisions Filipino philosophy to achieve. Yet Pada finds such grassroots philosophy as Mercado’s inadequate in reaching academic legitimacy unless “it begins to participate in the long tradition of discourse in philosophy.”<sup>48</sup> Discourse for Pada means an active engagement of one’s ideas with other traditions in philosophy. He regards the work of Florentino Hornedo as an example of a discursive philosophy: Hornedo’s book titled *The Power to Be* is a treatise about freedom that interacts with scholastic, phenomenological, and existential traditions that have their own

---

<sup>42</sup> Roland Theuas DS. Pada, “The Methodological Problems of Filipino Philosophy,” in *Kritike: An Online Journal of Philosophy*, 8:1 (2014), 24–44.

<sup>43</sup> *Ibid.*, 27.

<sup>44</sup> *Ibid.* Emphasis in the original.

<sup>45</sup> *Ibid.*, 28.

<sup>46</sup> *Ibid.*

<sup>47</sup> *Ibid.*, 36.

<sup>48</sup> *Ibid.*

distinctive view about the freedom and autonomy of persons.<sup>49</sup> Without a constant interaction with other schools of thoughts and philosophical traditions, the grassroots approach to Filipino philosophy seems for Pada philosophically inadequate and unlikely to bear much fruit.

Finally, Pada concludes by noting that “[t]he usefulness of philosophy as a tool for methodological understanding of culture is clear when it loses the character of simply establishing Filipino identity as a cultural edifice.”<sup>50</sup> This is in direct contrast to Mercado’s view of philosophy as a form of nationalism. In fact, Pada is explicit that the building of cultural identity is not a task for philosophers as Mercado assumed, since philosophy is a way of thinking that is enriched by constant intellectual interactions, both critical and constructive. Thus, attempting to merely describe how Filipinos actually think will always fall short of constituting an academically respectable philosophy. Pada ends by stressing that asking “whether philosophy is dead or is about to be born in our culture is irrelevant,” since what is needed is using these grassroots elements of Filipino philosophy, i.e., the data on how Filipinos actually think about certain philosophical topics, as starting point from which to build philosophical engagements with philosophical traditions so that eventually “our own discourse philosophy will prevail.”<sup>51</sup> In this regard, Pada clearly echoes the sentiments of Ferriols in the beginning of this paper. Thus, for Pada, Mercado and Timbreza’s version of Filipino philosophy, while valuable, is of limited philosophical applicability.

Indeed, there is much to agree with Pada’s assessment of Mercado’s approach to Filipino philosophy. It is in my view correct that if Mercado’s version of Filipino philosophy is purely descriptive, that is, merely seeking to describe what particular Filipinos think about philosophically-laden topics, then it is not even clear why it should count as *philosophizing*. In fact, a difficulty confronting Mercado is how exactly different is his work, if he considers it a philosophical work, with that of the social scientist. Pada’s call for the need for discourse in philosophy thus moves the conversation forward: a work is philosophical if it moves beyond purely describing people’s thoughts (as what Mercado had done) towards serious engagement, critical or constructive, with other intellectual traditions in philosophy. However, this need not be taken as a complete refutation of Mercado’s project but a need for expansion. In terms of methodology, one will benefit not only from linguistic analysis initiated by Mercado but also by controlled experiments and qualitative interviews as methods in experimental philosophy. Will the results of these studies remain descriptive and thus

---

<sup>49</sup> *Ibid.*, 39.

<sup>50</sup> *Ibid.*, 41.

<sup>51</sup> *Ibid.*, 43. Emphasis in the original.

inadequate to be academically respectable philosophy? The answer will depend on how the results will be employed. But if the results have uncovered a set of Filipino intuitions that are non-existent in the present literature or different from the Western ones that are taken as orthodox, it is interesting what philosophical implications would such intuitions have when analyzed to their logical end. And it is safe to say that exploring such philosophical implications is itself a worthy philosophical project. An exploration of certain philosophical intuitions—usually expressed in terms of principles, axioms, or commonsensical presuppositions—requires knowing the present intellectual terrain where their application would be relevant, and this presupposes a need for discourse where longstanding traditions, commonly used approaches, and venerable schools of thought can be utilized for intellectual engagement.

Not surprisingly, something like Pada's charge against Mercado has been a staple objection to experimental philosophy. After all, are experimental studies not merely describing what certain people think about a certain subject? And even if it happens that the intuitions of Filipino non-philosophers are contrary to that of the Filipino philosopher, is that supposed to change or sway the position of the Filipino philosopher who has expectedly spent a larger amount of time thinking and learning about the subject? Of course not, and experimental philosophers concur.<sup>52</sup> But this does not mean that the experimental results have no philosophical insights to offer. As Joshua Knobe and Shaun Nichols remarked:

The mere fact that a certain percentage of subjects hold a particular view cannot on its own have a significant impact on our philosophical work. Instead, it must be that the statistical information is somehow helping us to gain access to some other fact and that this other fact—whatever it turns out to be—is what is really playing a role in philosophical inquiry.<sup>53</sup>

Indeed, the possible differences of intuitions that may be unearthed are revelatory of whatever intellectual frameworks are at work between these two groups, and exploring the further implications of such frameworks are a philosophical task in itself. This is not so different from exploring the implications of worldviews that are opposed to each other, such as in the case of theism and atheism. For the most part theists and atheists disagree about the degree by which certain intuitions can persuade but these core

---

<sup>52</sup> Knobe and Nichols, "An Experimental Philosophy Manifesto," 6.

<sup>53</sup> *Ibid.*



disagreements need not prevent either of them to sift through what these intuitions may result to when applied to their logical conclusions. Consider for instance the project of atheist philosopher Erik Wielenberg when he explored the moral implications of an atheistic worldview or “a Godless universe” according to his book’s title.<sup>54</sup> One need not hold the same atheistic assumptions as Wielenberg to appreciate the value of his project: it can be a good way for theists to have a taste of the implications of a contrary worldview so they can weigh those considerations when assessing the overall reasonableness of their own. On the theist’s side, there is Christian philosopher Alvin Plantinga who, in his “Advice to Christian Philosophers,” has called for more wholeness and integrality within the Christian philosophical community. By this he means that a Christian philosopher, by virtue of being Christian, “has a perfect right to the point of view and pre-philosophical assumptions he brings to philosophic work; the fact that these are not widely shared outside the Christian or theistic community is interesting but fundamentally irrelevant.”<sup>55</sup> Bringing Plantinga’s advice, Christian philosophers ever since have traversed territories not typically explored by other philosophers solely because they constitute the core assumptions of the Christian faith: issues such as the Incarnation, the Trinity, Atonement and even the Eucharist. This means that bringing one’s philosophical skills to bear on topics that are rooted in fundamental, albeit not universally held, intuitions, is in itself a fruitful philosophical endeavor.

Can we benefit from Wielenberg’s project or Plantinga’s advice in terms of our search for a Filipino philosophy? I think we can, but unlike the Christian or the atheist worldview, the core assumptions embedded in the Filipino culture are still yet to be deeply explored. The good news is that we have good reason to believe that there are core assumptions in the Filipino philosophical mind that are vastly different or remain unknown to the Western one, as the pioneering works in experimental philosophy and Mercado’s own works have suggested. Thus, to say that the designation of “Filipino” in Filipino philosophy is just a matter of aesthetics<sup>56</sup> does not hold water at least from the perspective of experimental philosophy. This is because the categorization of “Filipino” is crucial to the kind of philosophy being argued for: it is philosophizing using the intuitions of Filipinos, especially those of the non-philosophers, as the building blocks of one’s argument or philosophical analysis.

---

<sup>54</sup> Erik Wielenberg, *Value and Virtue in a Godless Universe* (Cambridge: Cambridge University Press, 2005).

<sup>55</sup> Alvin Plantinga, “Advice to Christian Philosophers,” in *Faith and Philosophy*, 1:3 (July 1984), 256.

<sup>56</sup> I thank an anonymous reviewer for suggesting this.

By laying out the importance even the necessity of experimental philosophy as it relates to discovering forms of Filipino philosophy is of course not to imply that armchair philosophizing is already obsolete or unnecessary; on the contrary, experimental philosophy can supplement in various ways the results of our own intuition-based philosophizing. This is also not to say that a tradition of critical discourse would be lacking. What is simply novel in this approach is the openness to foundational intuitions as revealed by Filipino participants in future studies of experimental philosophy. These intuitions may then be used to provide the framework to engage in critical discourse with other schools of thought or pertinent philosophical debates. Finally, the enterprise is not redundant since even though the methodologies to be used are those of the empirical sciences, they are to be used in answering, clarifying, and addressing traditional philosophical questions, something that has not been thought possible before.<sup>57</sup>

Thus, with the exciting discovery that Filipino intuitions are worth looking at for their possible wide-ranging differences from Western ones, then it may already be the right time for Filipino philosophers to roll up their experimental sleeves and get to work.

*Humanities Department, Palawan State University, Philippines*

## References

- Abarbanell, Linda, and Marc D. Hauser, "Mayan Morality: An Exploration of Permissible Harms," in *Cognition*, 115 (2010).
- Ahlenius, Henrick and Torbjörn Tännsjö, "Chinese and Westerners Respond Differently to the Trolley Dilemmas," in *Journal of Cognition and Culture*, 12:3–4 (2012).
- Alexander, Joshua, *Experimental Philosophy: An Introduction* (Cambridge, UK: Polity Press, 2012).
- Andow, James, "Qualitative Tools and Experimental Philosophy," in *Philosophical Psychology*, 29:8 (2016).
- Batoon, Emmanuel D., "Tracing Mercado's Anthropological Perspective (Second of Two Parts)," in *Kritike: An Online Journal of Philosophy*, 8:2 (December 2014).
- Broderick, Damien, "Introduction II: Philosophy on the Inclined Plane," in *Philosophy's Future: The Problem of Philosophical Progress*, ed. by Russell Blackford and Damien Broderick (New York: Wiley-Blackwell, 2017).

---

<sup>57</sup> I thank an anonymous reviewer for raising the last two concerns.

- Buckwalter, Wesley and Stephen Stich, "Gender and Philosophical Intuition," in *Experimental Philosophy*, Vol. 2, ed. by Joshua Knobe and Shaun Nichols (New York: Oxford University Press, 2013).
- Costa, Albert, Alice Foucart, Sayuri Hayakawa, Melina Aparici, Jose Apesteguia, Joy Heafner, and Boaz Keysar, "Your Morals Depend on Language," in *PLoS ONE*, 9:4 (2014).
- De Leon, Emmanuel, and Marvin Einstein Mejaro, "An Interview with Leonardo Nieva Mercado, SVD," in *Kritike: An Online Journal of Philosophy*, 10:2 (December 2016).
- Deutsch, Max, "Gettier's Method," in *Advances in Experimental Philosophy and Philosophical Methodology*, ed. by Jennifer Nado (New York: Bloomsbury, 2016).
- \_\_\_\_\_, *The Myth of the Intuitive: Experimental Philosophy and Philosophical Method* (Cambridge: The MIT Press, 2015).
- Feltz, Adam and Edward T. Cokely, "Do Judgments about Freedom and Responsibility Depend on Who You Are? Personality Differences in Intuitions about Compatibilism and Incompatibilism," in *Consciousness and Cognition*, 18:1 (2009).
- Ferriols, Roque J., "A Memoir of Six Years," in *Philippine Studies*, 22 (1974).
- Fischer, Eugen and John Collins, "Introduction," in *Experimental Philosophy, Rationalism, and Naturalism: Rethinking Philosophical Method*, ed. by Eugen Fischer and John Collins (New York: Routledge, 2015).
- Henrich, Joseph, Steven J. Heine, and Ara Norenzayan, "The Weirdest People in the World?" in *Behavioral and Brain Sciences*, 33:2-3 (2010).
- Knobe, Joshua and Shaun Nichols, "An Experimental Philosophy Manifesto," in *Experimental Philosophy*, Vol. 1, ed. by Joshua Knobe and Shaun Nichols (New York: Oxford University Press, 2008).
- Machery, Edouard, "The Illusion of Expertise," in *Experimental Philosophy, Rationalism, and Naturalism: Rethinking Philosophical Method*, ed. by Eugen Fischer and John Collins (New York: Routledge, 2015).
- Machery, Edouard, Christopher Y. Olivola, and Molly De Blanc, "Linguistic and Metalinguistic Intuitions in the Philosophy of Language," in *Analysis*, 69 (2009).
- Machery, Edouard, Ron Mallon, Shaun Nichols, and Stephen Stich, "Semantics, Cross-Cultural Style," in *Cognition*, 92:3 (2004).
- Markus, Hazel and Shinobu Kitayama, "Culture and the Self," in *Psychological Review* 98 (1991).
- Mercado, Leonardo, *Elements of Filipino Philosophy* (Tacloban City: Divine Word University Publication, 1974).
- \_\_\_\_\_, "Filipino Thought," in *Philippine Studies*, 20:2 (1972).
- \_\_\_\_\_, "Reasoning," in *The Filipino Mind* (Washington, D.C.: The Council for Research in Values and Philosophy, 1994).

- \_\_\_\_\_, "Reflections on the Status of Filipino Philosophy," in *Kritike: An Online Journal of Philosophy* 10:2 (2016).
- \_\_\_\_\_, "What is Philosophy?" in *Filipino Thought* (Manila: Logos Publications, Inc., 2000).
- Mortensen, Kaija and Jennifer Nagel, "Armchair-Friendly Experimental Philosophy," in *A Companion to Experimental Philosophy*, ed. by Justin Sytsma and Wesley Buckwalter (Oxford: Blackwell, 2016).
- Nisbett, Richard E., *The Geography of Thought: How Asians and Westerners Think Differently ... and Why* (New York: The Free Press, 2003).
- O'Neill, Elizabeth and Edouard Machery, "Experimental Philosophy: What is it Good For?" in *Current Controversies in Experimental Philosophy*, ed. by Edouard Machery (New York: Routledge, 2014).
- Pada, Roland Theuas DS., "The Methodological Problems of Filipino Philosophy," in *Kritike: An Online Journal of Philosophy*, 8:1 (2014).
- Plantinga, Alvin, "Advice to Christian Philosophers," in *Faith and Philosophy*, 1:3 (July 1984).
- Rose, David and David Anks, "In Defense of a Broad Conception of Experimental Philosophy," in *Metaphilosophy* 44 (2013).
- Silverman, Jason, Elizabeth Hall, Jamie Aten, Laura Shannonhouse, and Jason McMartin, "Christian Lay Theodicy and the Cancer Experience," in *Journal of Analytic Theology* 8 (2020).
- Stich, Stephen and Kevin P. Tobia, "Experimental Philosophy and the Philosophical Tradition," in *A Companion to Experimental Philosophy*, ed. by Justin Sytsma and Wesley Buckwalter (Oxford: Blackwell, 2016).
- Stuart, Michael T., Yiftach Fehige, and James Robert Brown, *The Routledge Companion to Thought Experiments* (New York: Routledge, 2018).
- Weinberg, Jonathan M., "Going Positive by Going Negative: On Keeping X-Phi Relevant and Dangerous," in *A Companion to Experimental Philosophy*, ed. by Justin Sytsma and Wesley Buckwalter (Oxford: Blackwell, 2016).
- Weinberg, Jonathan, Shaun Nichols, and Stephen Stich, "Normativity and Epistemic Intuitions," *Philosophical Topics*, 29:1/2 (2001).
- Wielenberg, Erik, *Value and Virtue in a Godless Universe* (Cambridge: Cambridge University Press, 2005).

Article

## Saint Paul: A Revolutionary Subject

---

*Delfo C. Canceran*

**Abstract:** This paper highlights Badiouian philosophy of the subject by making use of the model of Paulinian vocation. The Badiouian subject is not defined prior to its vocation but is made a subject by the call to revolutionary movement. In the Paulinian model, Paul is called by the risen Jesus, and Paul pursued that call by building a community defined by faith and not by law. The universal community is marked by indifference to the differences of its members. The Badiouian contribution to the social movement is clearly the presentation of the vocation of Saint Paul in building a new humanity based on equality and inclusiveness.

**Keywords:** Badiou, Event, subject, truth

*I read Paul as a text about a new and provocative conception of truth and, more profoundly, about the general conditions for a new truth... I read Paul as a human creation in the field of the question of truth.<sup>1</sup>*  
— Alain Badiou

Badiou is considered as a philosopher of the new.<sup>2</sup> However, some skeptics ask if there is something new under the sun? There seems to be preponderance to the old and the same expressed in this skepticism. Of course, we are not speaking of literal things under the sun, but this suggests that the sun exposes everything to visibility and clarity so that everything becomes evident in our sight or vision. However, in this world where not everything is clear and obvious, we need to inquire into it or wonder about it. Moreover, conservatives suspect the new because it upsets the things as usual or the order of things in the world. Tradition, as the reproduction of the same, seems to be the natural and sacrosanct thing, thus,

---

<sup>1</sup> Adam S. Miller and Alain Badiou, "Universal Truths and the Question of Religion: An Interview with Alain Badiou," in *Journal of Philosophy and Scripture*, 3:1 (2005).

<sup>2</sup> See Ed Pluth, *Badiou: A Philosophy of the New*, (Cambridge: Polity Press, 2010).

the new seems aberrant or fake. Unable to handle the new, the conservatives discard or discount it, and stick with or adhere to the old tradition for complacency and security.<sup>3</sup>

To demonstrate his penchant of the new, Badiou writes about the life of Paul, not as a theological piece or a hermeneutical treatise that endeavors to interpret his letters, but a philosophical and a revolutionary enterprise that concretizes this new in the event: the resurrection of Jesus.<sup>4</sup> For Badiou, Paul exemplifies the militant or activist figure of the event that enables the latter to embark into a new endeavor and spread Christianity. Paul demonstrates the faithful subject to the event. This Paulinian model of revolutionary spirit is badly needed in this time of dearth of leaders and fatigue to activism in some militant movements. Badiou shows that the conversion or vocation of Paul leads him to embrace militant engagement in the world. This engagement requires fidelity to the event that enables him to persevere or persist. This new experience of Paul makes him a relevant figure of revolutionary movement.<sup>5</sup>

### Vision of Saint Paul

Badiou presents the life of Paul, the Apostle as recorded mostly in his biblical writings, to exemplify his understanding of the event and activism that provides a revolutionary figure in our history.<sup>6</sup> In the interview, Badiou states that he reads Paul not as a sacred text or a hermeneutic task, but as a political figure with revolutionary passion. Moved by that event in a vision, Paul is impelled by a new truth that questions and transforms the world. In particular, Paul's conversion or vocation transformed him into a new subject of revolutionary movement in the making of the church. Paul consistently proclaims the resurrection of Jesus Christ as the inspiration of this new movement. Paul began as a persecutor of the people who followed the way of Jesus Christ but in his Damascus experience, he was altered. That new experience compelled him to pursue a different path and enter into a missionary work. In the Acts of the Apostles, it is written:

---

<sup>3</sup> See Thomas S. Kuhn, *The Structure of Scientific Revolution* (Chicago: Chicago University Press, 1962).

<sup>4</sup> Alain Badiou, *Saint Paul: The Foundation of Universalism*, trans. by Ray Brassier (Stanford: Stanford University Press, 2003).

<sup>5</sup> There is a question on the correct description of Paul's experience in Damascus when the risen Jesus Christ appeared and conversed to him. For some, this particular experience is not a conversion but a vocation because he is called by Jesus Christ to become an apostle and member of the church. See Stanley E. Porter, "Paul as Jew, Greek and Roman: An Introduction," in *Paul: Jew, Greek and Roman*, ed. by Stanley E. Porter, (Leiden: Brill, 2008).

<sup>6</sup> Badiou, *Saint Paul*, 2.

Now as he [Saul/Paul] was going along and approaching Damascus, suddenly a light from heaven flashed around him. He fell to the ground and heard a voice saying to him, "Saul, Saul, why do you persecute me?" He asked, "Who are you, Lord?" The reply came, "I am Jesus, whom you are persecuting. But get up and enter the city, and you will be told what you are to do." The men who were traveling with him stood speechless because they heard the voice but saw no one. Saul got up from the ground, and though his eyes were open, he could see nothing; so, they led him by the hand and brought him into Damascus. For three days he was without sight, and neither ate nor drank (Acts 9:4-9 NRSV).

Paul was bent on chasing the Christians who, when arrested, were ruthlessly persecuted. He persecuted them because they proclaimed that the Messiah had already come and, being a Pharisee himself, this proclamation threatened the authority of the law.<sup>7</sup> In that passage, Paul was planning to hunt more Christians in Jerusalem. But on his way to Damascus, a light flashed on him and blinded him. That encounter between stunned Paul and the risen Jesus opened a new present different from his previous life. He was seized by that vision that propelled him to embark on a new path. This visionary encounter with the risen Jesus called him to a mission. Ordinarily, conversion is a long and slow process of transformation. In this particular case, this encounter seems to suggest that Paul was converted easily, and led to a new life. We may say that, by the grace of God, he was quickly converted. In his letters, grace was prominent in his teachings, relentlessly defending it against finding favor by mere human effort and achievement.<sup>8</sup>

However, we can also say that he can easily shift to being a Christian due to his hybrid identity: belonged to a Diaspora Jew, gained Greek education, and claimed Roman citizenship rolled into a unique personality. These three worlds—Jewish, Greek, and Roman—provided him an advantage in his preparation and knowledge, and set him to embrace his new vocation or call becoming the apostle to the Gentiles. Thus, this multiple identity allowed him to be receptive and open to new experience and fresh venture. For example, being a Pharisee, he believed in the resurrection of the dead and, when the risen Jesus appeared to him, he was easily convinced of

---

<sup>7</sup> Jerome Murphy-O'Connor, *Paul: His Story* (Oxford: Oxford University Press, 2004), 23.

<sup>8</sup> Klaus Hacker, "Paul's Life," in *The Cambridge Companion to Saint Paul*, ed. by James D.G. Dunn (Cambridge: Cambridge University Press, 2003), 24.



that vision. That identity facilitated his shift to being a Christian by proclaiming the resurrected Jesus.<sup>9</sup>

After his conversion, Paul has embraced a new way of life. He takes the radical view of faith in fidelity to that event. This shift shows the break in the life of Paul, first as a persecutor of Christians then, as apostle to the Gentiles. This break is laden with mystery. Due to his preference for the new, Badiou subscribes to the historical break since he posits the necessity of an event that rips or ruptures the continuity of life. That event must be very significant because it impacts the life of the person. Thus, the model of Paul answers the search for a radical figure worth emulating in our own time. Paul's vision on the road to Damascus equipped him with an entirely new perspective, deeply rooted in that encounter with the risen Jesus. This new perspective summons him to be an apostle by proclaiming the resurrection, which inaugurates a new world where justice and peace reign among the people.<sup>10</sup>

### Event of Truth

When Badiou speaks of the event, he does not refer to the flat and typical happening around us, but rather, to the singular and transformative episode that befalls us in the world.<sup>11</sup> An event is not just an occurrence that befalls a person, but an experience that transforms the person. Event is the sudden irruption of the new the effect of which is compelling and enduring to the person. The event that happened in Damascus transformed the life of Paul. This event cannot be captured by the available knowledge but, seeks a new language that articulates it.<sup>12</sup> Moreover, that event brings forth truth by transforming the world. The truth interrupts or breaks the world as usual. Equipped with this truth, Paul persistently challenges the order of things. In fact, this event is a driving force that alters the world. By contesting the reigning present, he insists on this new order that should replace and govern the new-found community. This new order crumbles the wall that separates the members of the community.<sup>13</sup> In his letter to the Galatians, Paul spoke of an inclusive community that overcomes division.

---

<sup>9</sup> See Richard Wallace and Wynne Williams, *The Three Worlds of Paul of Tarsus* (London: Routledge, 1998).

<sup>10</sup> N.T. Wright, *What Saint Paul Really Said: Was Paul of Tarsus the Real Founder of Christianity?* (Michigan: William B. Eerdmans Publishing Company, 1997), 37.

<sup>11</sup> Badiou provides the example of this event in the historic May 1968 revolution in France that significantly changed various institutions in the country like the workplaces, universities, arts, and so on. See Alain Badiou, *The Rebirth of History: Times of Riots and Uprisings*, trans. by Gregory Elliott, (London: Verso, 2012).

<sup>12</sup> Christopher Norris, *Badiou's Being and Event* (London: Continuum, 2009), 9.

<sup>13</sup> Badiou, *Saint Paul*, 108.

As many of you as were baptized into Christ have clothed yourselves with Christ. There is no longer Jew or Greek, there is no longer slave or free, there is no longer male and female; for all of you are one in Christ Jesus (Galatians 3:26-28 NRSV).

The event of the resurrection of Jesus surmounts or obliterates all divisions such as ethnicity, status or gender. Faith in Christ is the sole requirement in this new community of believers. Badiou carefully establishes the singularity of Paul's position in the controversy of difference and opposition between the Jewish ethnic election (and law), and the Greek love of logic (philosophy). The event neither needs selection nor logic because it overcomes or transgresses it. In Christ, through baptism, the difference or opposition of traditions that mark people off in the community disappears because there is a new standard established by faith.<sup>14</sup> The event challenges this status that sets the Jews and the Greeks apart. In the confrontation between Paul and Peter on matters of common meals, we see Paul opposing Peter separating himself from the common meal, and dividing the community composed of Jewish Christians, and Gentile Christians.

But when Cephas [Peter] came to Antioch, I opposed him to his face, because he stood self-condemned; for until certain people came from James, he used to eat with the Gentiles. But after they came, he drew back and kept himself separate for fear of the circumcision faction. And the other Jews joined him in this hypocrisy, so that even Barnabas was led astray by their hypocrisy. But when I saw that they were not acting consistently with the truth of the gospel, I said to Cephas before them all, 'If you, though a Jew, live like a Gentile and not like a Jew, how can you compel the Gentiles to live like Jews?' (Galatians 2:11-14 NRSV).

Paul is cognizant of the community consisting of Jews and Greeks. He sticks to the new order by respecting their differences without imposing any privileged practice. He recognizes and accepts these differences that mark the members. This passage diverges from the amicable agreement forged in Jerusalem and leads to a confrontation between them. For Paul, Peter's separation violated the agreement of the unity of the community. By

---

<sup>14</sup> Daniel Boyarin, *A Radical Jew: Paul and the Politics of Identity* (Berkeley: University of California Press, 1994), p. 16.

Peter's gesture, the Gentiles are rejected from table fellowship with the Jews and, by implication, to be admitted to the table, the Gentiles need to renounce their Greek custom, and by force, assume the Jewish custom. That confrontation on table fellowship shows that Paul is persistent on the demand for inclusion that would govern the community, and overturns that division. The new agreement of respect and inclusion in the emerging community exceeds the evident difference between the Jews and the Greeks. The difference between the two is superseded by a new truth by transcending such difference. This transcendence does not demolish their customs and practices but insists on an indifference to this difference. Truth procedure can only offer indifference to the differences that exist between the Jews and the Gentiles in forming an inclusive church.

In Badiou's view, the event is closely linked with truth. Badiou re-conceptualizes truth neither as a correspondence fitted to external reality nor as a consensus established by a group. Basically, truth is a process of making a particular set of statements or discourses which start from the event and, from that starting point, follows from it. In short, the event precedes truth. The setting off is important since it signals a break from the continuity created by that singular event. Moreover, the event creates a new possibility, and inaugurates a new situation. In effect, the event alters the state of affairs. The truth comes after the event, thus, a truth is always the truth of the event that brings about a new situation. The conversion experience of Paul is a truth of that particular encounter. It is a subjective truth experienced by Paul who is summoned to start off a new mission. This truth cannot be captured by the state of affairs or the regime of knowledge that defines or confines the event to a repetition or reproduction of the same. Badiou usually uses the phrase "truth procedure" to drive his point that truth does not descend from above, or emanate from below. Truth is an ongoing practice and struggle. The event creates truth, and this truth is a procedure because it is an ongoing creation of the new situation. However, this truth procedure is not reducible to this process or its outcome. Truth always exceeds or surpasses the process or outcome because it continuously provides a driving force or desire that unfolds in history. For Badiou, Paul's experience of that singular event of the risen Jesus should be described as procedure rather than as content. The subjective "event" is an impetus for his persistence in the creation of a new community continuing and expanding into universalism.

Moreover, truth is intimately connected with agency. Truth is not about contemplation or discernment but about intervention or resistance in the world. One does not simply know or learn a truth; one performs it as an agent or actor that changes the status. For Badiou, the community of believers is creating a new truth of inclusion or universality, where Jews and Greeks are no longer divided, but respected. This community is an agent of change

for inclusion and transformation. Moreover, this agency is always a collective effort, not individual endeavor. Although this is a collective endeavor, agency is not confined to a particular group having shared characteristics, such as the Jews or the Greeks, but transforms the community into an inclusive or universal openness to differences of the believers. Badiou insists that truth is not possessed by any particular group, status, and privilege but rather, shared by the whole community. This universalism is indifferent to the differences of people.

### Subject of Truth

The subject does not refer to any individual or all social actors, but rather, refers to the actor who acts on the event. The subject does not align with the existing state or status, but radically alters the old order and creates a new situation. To effect change, the subject intervenes in the world. Thus, the subject effectuates a break or cut, and inaugurates or broaches a new situation.<sup>15</sup> As a caveat, Badiou does not define the actor as a cause, but rather as the *effect* of an event. The subject can only exist after making a critical decision; the subject is produced through this decisive action. The subject does not choose a truth but rather truth creates the subject. The subject does not precede a truth but is constituted or produced by the truth. Thus, the subject “stands as evidence of the presence of something new in the situation.”<sup>16</sup> In the case of Paul, the Damascus event produces a truth that compels him to decide and to act accordingly.

According to Badiou, Paul demonstrates the challenge to politics and the critique of the law. Badiou connects the making of the subject, and the questioning of the law. The subject is constituted no longer by the law, but by faith. Thus, faith is characterized by freedom from the law. Unbounded from the law, faith exceeds or surpasses the law. Gentiles are no longer in anyway different from the Jews because the Christ or Messiah is not just exclusively possessed by the Jews but given to the whole community who professes the faith in the risen Jesus. Thus, Badiou disconnects the subject and the law and connects the subject and faith.<sup>17</sup> In his Letter to the Galatians, Paul asserts:

We ourselves are Jews by birth and not Gentile sinners; yet we know that a person is justified not by the works of the law but through faith in Jesus Christ. And

---

<sup>15</sup> Bruno Besana, “The Subject,” in *Alain Badiou: Key Concepts*, eds. A.J. Barlett & Justin Clemens (Durham: Acumen, 2010), 40.

<sup>16</sup> Pluth, *A Philosophy of the New*, 104.

<sup>17</sup> Stanislas Breton, *A Radical Philosophy of Saint Paul* (New York: Columbia University Press, 2011), 92–93.

we have come to believe in Christ Jesus, so that we might be justified by faith in Christ, and not by doing the works of the law, because no one will be justified by the works of the law (Gal. 2:15-16 NRSV).

In this new order, the subject is no longer enslaved by the law that perpetuates the status or repeats the same, but freed from the law because the subject is inspired by faith in the risen Jesus. It is no longer law that rules but faith in Jesus Christ. It is no longer obedience to the law that matters, but faith in the risen Jesus. By the event of the resurrection, the subject is driven by faith that justifies; not the law that enslaves. The event of the resurrection produces a truth that the subject lives on to create a new situation. The truth unfolds in the process and creates new situation. The event calls a subject to a militant or activist vocation. That event demands fidelity of the subject. This fidelity is not an automatic repetition of the same but radical transformation of the order of things or the things as usual. The event calls for the creation of new situation. The decision made is not derived from the state or status but based only in faithfulness rooted in the event. The subject keeps on exploring the possibilities of the event; the possibilities are infinite or numerous. The subject does this exploration by mapping a new element in the situation, and by taking elements from different parts of the situation and forming them into a new set.

Moreover, the event requires persistence in the quest for truth. Badiou demands that a subject keeps on moving, regardless of the consequences involved in that action. However, this is not simply a matter of following one's own path, but a decision compelling one's own faith in pursuing a new path. The subject has to take the risks and stakes life. A subject must be indifferent to what others think around him and clings on to the event leading him/her to try a new trail. Thus, the subject is a revolutionary actor in the sense that it is consumed by the event that propels him/her to intervene in the world or resist the regime of things. An event has to create a new language for a new situation. New words must be invented, employed, and expressed beyond or outside the usual order. This new language is not prearranged or ready-made but an ordeal in the search for language. It is through a truth procedure that this new language is elaborated and specified in the process.

A truth procedure must be read on its own terms, not on other criteria derived from the past or given by a group. A truth procedure is a process based on the singularity of the event, and not on the hegemony of knowledge. It unfolds in an unpredictable manner, depending on the commitment and the movement of its subjects who pursue the truth of the event. The procedure concludes with a transformation of the state into a new situation. The

situation is redefined in view of the truth envisioned by the subject. The entire state has to be rearranged to make room for the formerly excluded part to be included into the new set. Reality must be transformed by the subject that creates a new situation using new language. New language is created for defining the situation.

### **Universalism in the Community**

Truth, as we have seen, is open to the possibility of universalism. Paul exemplifies that possibility of universalism when he insists on the universalizing message of the resurrection to the people by freeing the community from the rigid enclosure of Judaism (cultural privileges), and conformity to the Roman law (political and legal strictures), and the philosophy of Greece (logic and rationality). Universalism is a process of openness unfastened by the event of the resurrection that alters the existing regime of division and hierarchy. To create something universal is to go beyond the existing order or norm. The event unfolds a new situation based on the truth of the event. Paul offers a new situation in the community that breaks from the previous setup. The community envisioned by Paul is a militant singularity inspired by the event of the resurrection, and operating under a radical universalism.

Paul is aware of the composition of the community. However, the new truth exceeds the evident difference that exists between these ethnic groups. We can only receive a new truth by going beyond such differences. This new situation does not mean that the people need to renounce their customs and practices. Instead, they become indifferent to the differences of one another so that they can build a new creation or new humanity. With this in mind, Paul seeks to reorient the members, not in relation to cultural specificities (ethnicity, status or gender), but in relation to truth. The Letter to the Ephesians spells it out:

But now in Christ Jesus you who once were far off have been brought near by the blood of Christ. For he is our peace; in his flesh he has made both groups into one and has broken down the dividing wall, that is, the hostility between us. He has abolished the law with its commandments and ordinances, that he might create in himself one new humanity in place of the two, thus making peace, and might reconcile both groups to God in one body through the cross, thus putting to death that hostility through it (Eph. 2:13-16 NRSV).

Paul is persistent with his vision of the Kingdom of God which demands fidelity to the Gospel. This vision encompasses the inclusive community where the division that separates the Jews and the Gentiles disappear and where accord among the members prevails. Thus, the community is characterized by harmony and reconciliation—not by hostility and conflict—guided by their faith in the risen Jesus. The resurrection has freed them from the bondage of the law and the faith in the risen Jesus provides freedom to the community. The law of the Jews inscribed in the Mosaic order is no longer binding but only the faith given by the grace of God rules the new humanity in the Kingdom of God.<sup>18</sup>

For Badiou, “something is universal if it is something that is beyond established differences [and] that these differences become indifferent.”<sup>19</sup> Thus, universalism is a creation of the event whose truth is open to prospective members who believe in the risen Jesus. Faith matters in the emerging community. “Universalism is always the result of a great process that opens with an event. To create something universal is to go beyond evident differences and separations.”<sup>20</sup> With this new truth, the community becomes indifferent to these differences. The Gospel brings about universalism where believers or followers participate in its realization. The resurrection is a singular event that proclaims the universalism of the community since it rejects the restrictions of the law or culture. Thus, Badiou refuses to follow the regimes of Jews and the Greeks in favor of universalism by becoming indifferent to differences of people.

### Problematic (In)Difference

Badiou criticizes the philosophers of difference such as Emmanuel Levinas and Jacques Derrida because they recognize the difference or defend the alterity of the other.<sup>21</sup> In that affirmation, the other essentially retains inequality or hierarchy since the other is given the upper hand in its relationship with the self. This inequality is understandable since the other has been historically wronged or oppressed by the self. These differences cannot be abrogated but must be recognized—such as sexual difference, ethnic difference, and colonial difference of peoples—because they interrogate the power that marginalizes or excludes them from society. The

---

<sup>18</sup> Timothy G. Gombis, *Paul: A Guide to the Perplexed* (London: Bloomsbury, 2010), 133–145.

<sup>19</sup> Miller and Badiou, “Universal Truths and the Question of Religion: An Interview with Alain Badiou.”

<sup>20</sup> *Ibid.*

<sup>21</sup> See Emmanuel Levinas, *Outside the Subject*, trans. by Michael Smith (California: Stanford University Press, 1994) and Jacques Derrida, *The Gift of Death*, trans. by David Wills (Chicago: Chicago University Press, 1995).



hegemonic power suppressed or repressed these differences so that they are hidden and pushed. These differences will always remain such that the other will continually hunt the self to account for these atrocities.

Badiou opts for the indifference to differences because of his overarching insistence on equality and inclusion of the people in the community, regardless of their differences in class, ethnicity, and gender, to name just a few. The community should accept individuals who profess faith in the risen Jesus and not on his/her ethnicity, gender, and class. Thus, some questions occur: What shall we do with these different marks of people who belong to the community? Do we just ignore them or do we need to address them, not just for the sake of inclusion or equality but also for the sake of justice and right? In that universal unity of the people, the differences are compromised. However, a compromise only temporarily hides the disagreement and inhibits the hostility that exists between two opposing groups. Moreover, tolerance implies that the dominant group allows a practice or action within their own standards. When those limits are transgressed, the dominant group will impose restraints to the unruly subordinate group.<sup>22</sup>

In particular, Derrida speaks about justice in contrast to the law. While the law is general in its application by treating the cases in the same way (The law applies to all), justice is singular in its application to different cases of people. Thus, the law highlights the generality of the same, and justice emphasizes the singularity of the other. The law relies on past decisions and applies to similar cases, while justice considers the past and the present. In the singularity of the other, the judge undergoes the ordeals of decision by detailing the situation of the other. There is always the irreducibility of the new situation different from the old situation. In effect, the decision blends the old and the new in the singularity of the other. The singularity of the other applies to woman in sexual differences, to colonies in racial discrimination, to blacks in racism, to gays in heterosexism, to migrants in globalization. The point of justice is to highlight the marginalization and exclusion of the other, and the interrogation and transformation of the situation. The democracy to come is the breaking of the same, and the opening of the space for the unexpected other, undefined by the past and even by the present. The other is always to come.<sup>23</sup>

The social teaching of the Church insists on the preferential love or option for the poor since, historically and politically, they have been deprived

---

<sup>22</sup> See Steve Clarke, Russell Powell & Julian Savulescu, *Religion, Intolerance and Conflict: A Scientific and Conceptual Investigation* (Oxford: Oxford University Press, 2013).

<sup>23</sup> Jacques Derrida, "Force of Law: The 'Mystical Foundation of Authority,'" trans. by Mary Quaintance in *Cardozo Law Review*, 11:5-6 (1990).

by society and subordinated to the elite. That unjust status should be rectified ,and addressed. The option for the poor is a commitment that demands social justice so that the poor can also benefit from the creation of God, and enjoy their lives in the community. There is no peace without justice. Peace is not just the negation of violence, but the positivity of life that endeavors for happiness and fulfillment. We need to address the social problems of people, and rectify injustices so that we create a just society, and benefit from the bounty of the earth. Thus, we need to be inclusive and fair but, as a requirement, we need social justice for those who have been dispossessed of the fullness of life since they suffer from dearth and bigotry. Thus, the priority is the building of the Church of the poor based on the imperative of the justice and the flourishing of life.

### Conclusion

Paul's encounter with the risen Jesus on the road to Damascus transformed his identity and propelled his proclamation of this singular event. Badiou claims that Paul recognizes that the resurrection of Jesus Christ is not pointing to some Jewish sign or Greek wisdom that could be bound to a particular group or fixed to a particular culture. That encounter is a singular and pure event, devoid of such exclusivities or particularities, and opens the path to universalism and equality. The event is indifferent to the differences of people because it is not tied to any group, and not restricted by any law; it is open to everyone willing to embrace that event and to perform the truth. This event is transformative because it inspires the subject to become a new creation or new humanity rooted in the resurrection. Paul is an apostle of revolutionary project where he forms community as a militant collection or multitude of subjects operating under an avant-garde movement toward a radical universalism. However, a problem arises on the minimalization of differences and the maximization of similarities of the people in view of inclusion and equality in the community which may undermine the struggles of the others who have suffered from discriminations and injustices.

*Ecclesiastical Faculties, University of Santo Tomas, Philippines*

### References

- Badiou, Alain, *Saint Paul: The Foundation of Universalism*, trans. by Ray Brassier, (Stanford: Stanford University Press, 2003).  
 \_\_\_\_\_, *The Rebirth of History: Times of Riots and Uprisings*, trans. by Gregory Elliott (London: Verso, 2012).

- Besana, Bruno, "The Subject," in *Alain Badiou: Key Concepts*, ed. by A.J. Barlett & Justin Clemens (Durham: Acumen, 2010).
- Boyarin, Daniel, *A Radical Jew: Paul and the Politics of Identity* (Berkeley: University of California Press, 1994).
- Breton, Stanislas, *A Radical Philosophy of Saint Paul* (New York: Columbia University Press, 2011).
- Clarke, Steve, Russell Powell & Julian Savulescu, *Religion, Intolerance and Conflict: A Scientific and Conceptual Investigation* (Oxford: Oxford University Press, 2013).
- Derrida, Jacques, "Force of Law: The 'Mystical Foundation of Authority,'" trans. by Mary Quaintance, in *Cardozo Law Review*, 11:5–6, (1990).
- \_\_\_\_\_, *The Gift of Death*, trans. by David Wills (Chicago: Chicago University Press, 1995).
- Hacker, Klaus, "Paul's Life," in *The Cambridge Companion to Saint Paul*, ed. by James D.G Dunn (Cambridge: Cambridge University Press, 2003).
- Kuhn, Thomas, *The Structure of Scientific Revolution* (Chicago: Chicago University Press, 1962).
- Levinas, Emmanuel, *Outside the Subject*, trans. by Michael Smith (California: Stanford University Press, 1994).
- Miller, Adam and Alain Badiou, "Universal Truths and the Question of Religion: An Interview with Alain Badiou," in *Journal of Philosophy and Scripture*, 3:1 (2005).
- Murphy-O'Connor, Jerome, *Paul: His Story* (Oxford: Oxford University Press, 2004).
- Norris, Christopher, *Badiou's Being and Event* (London: Continuum, 2009).
- Pluth, Ed, *Badiou: A Philosophy of the New*, (Cambridge: Polity Press, 2010).
- Porter, Stanley, "Paul as Jew, Greek and Roman: An Introduction," in *Paul: Jew, Greek and Roman*, ed. by Stanley E. Porter (Leiden: Brill, 2008).
- Timothy Gombis, *Paul: A Guide to the Perplexed* (London: Bloomsbury, 2010).
- Wallace, Richard, and Wynne Williams, *The Three Worlds of Paul of Tarsus*, (London: Routledge, 1998).
- Wright, N.T, *What Saint Paul Really Said: Was Paul of Tarsus the Real Founder of Christianity?*, (Michigan: William B. Eerdmans Publishing Company, 1997).

## **Midgley, Mary, *What Is Philosophy For?*<sup>1</sup>**

---

**Jovito V. Cariño**

The book *What Is Philosophy For?* is Mary Midgley's final published work. It came out in 2018, the same year she passed on, or just a year shy of what would have been her 100<sup>th</sup> birthday. She was born in London on September 13, 1919. Midgley was, for me, a belated but precious find. I had the distinct privilege of meeting and listening to her in person when I attended the International Society for MacIntyrean Enquiry (ISME) held at the University of Nottingham, UK sometime July 2012. Prior to that, Midgley was not even a name to me (due in large measure to the limited reach of my philosophic reading). Like any average researcher, I used to think of English philosophy or any English philosopher for that matter as either analytic or masculine (logic, mathematics, and analysis are traditionally recognized as male provinces of the human brain). My naiveté had me consider either as a rarity or an oddity an English intellectual who forays into Continental philosophic questions. Obviously, I failed to reckon that the likes of R.G. Collingwood, Terry Eagleton, and Simon Critchley were English and so were woman thinkers like Iris Murdoch, Philippa Foot, and Elizabeth Anscombe who were not just as compelling but also as intellectually gifted. Midgley belonged to this strand of English thought, a tradition that stretched as far back to Lord Shaftesbury, David Hume, and, to a degree, Adam Smith of *The Theory of Moral Sentiments*. She counted herself as one among those who were sympathetic with the so-called "big questions" frowned upon by the analytic movement. Simon Critchley, for example, reported how Jacques Derrida almost missed his honorary doctoral degree from the University of Cambridge after its academic community questioned whether there was genuine philosophy in Derrida's body of works.<sup>2</sup>

As her culminating work, *What Is Philosophy For?* weaves together the main themes Midgley explored and navigated in her previous projects. In a way, with or without her conscious intention, it provides a general tour of her philosophic itinerary spanning a lifetime of intellectual labor from *Beast and*

---

<sup>1</sup> London: Bloomsbury Academic, 2018, 223pp.

<sup>2</sup> Simon Critchley, *Continental Philosophy: A Very Short Introduction* (Oxford: Oxford University Press, 2001), 34.

*Man* (1978), *Animals and Why They Matter* (1983), *Evolution as Religion* (1985) and *Wisdom, Information, and Wonder: What Is Knowledge For?* (1989), to the more recent vintage represented by the likes of *Science as Salvation: A Modern Myth and Its Meaning* (1992), *The Ethical Primate: Humans, Freedom, and Morality* (1994), and *Utopias, Dolphins, and Computer* (1996). Consequently, what veritably counts as her swan song also serves as an overture which highlights the contours of her fundamental claims on such problems touching on human nature, ethics, freedom, scientism, materialism, quantum physics, digital technology, and yes, the function of philosophy, among others. The collection of topics covered by the book may appear daunting but due to Midgley's lucid, almost crystalline prose, not to mention her characteristic natural way with words, reading it easily becomes a refreshing experience of philosophic reeducation. In describing what she thought of philosophy, for example, she wrote: "...I have often suggested that philosophy is best understood as a form of plumbing. It's the way in which we service the deep infrastructure of our lives—the patterns that are taken for granted because they have not really been questioned."<sup>3</sup> In another part, where she dismissed the reductive physicalism of the Vienna Circle and Rudolf Carnap, she argued her point by saying: "If 'physical entities' means only ones that can be described in the language of physics, then everyday life simply contains hardly any physical entities at all. Physics never speaks of loaves and apples, pen and paper, men and women, bricks and mortar. It always speaks, far more abstractly, of solids and liquids, protons and electrons, vacuums and black holes."<sup>4</sup>

Midgley's notion of philosophy is guided by her agonistic ontological vision which puts her in direct opposition to the segmentary proclivity of the modern worldviews. She attributes the spread of the latter to the "increasing specialization of our age—the growing tendency of educators to supply more and more separate examinable qualifications for everything rather than putting things together intelligibly."<sup>5</sup> While she does not discount the plurality of perspectives surrounding a problem, she believes nonetheless that they are not completely estranged so as not to find any connection among them. Hence, for her, philosophy's task is "...to find ways of bringing the two sides together."<sup>6</sup> More than just an epistemic stance, this intellectual attitude also suggests Midgley's overall ethical view. This is shown in her critique of the Vienna Circle's physicalism and BF Skinner's behaviorism. Against the latter theories, Midgley argues against the lopsided focus of those theories on

---

<sup>3</sup> Mary Midgley, *What Is Philosophy For?* (London: Bloomsbury Academic, 2018), 64.

<sup>4</sup> *Ibid.*, 152–153.

<sup>5</sup> *Ibid.*, 192.

<sup>6</sup> *Ibid.*, 194.

## 112 WHAT IS PHILOSOPHY FOR?

the outside dimension of the human lifeworld at the expense of the inner subjective spirit. As she wrote, "...the inner, subjective point of view is every bit as natural and necessary for human thought as the outside objective one...you can't have the outside of the teapot without the inside."<sup>7</sup>

Given the obvious slant of the title, the whole trajectory of the book comes to the reader almost gratuitously. In choosing to formulate the title in this manner, Midgley runs the risk of getting her portrait of philosophy read as functional, utilitarian, or even apologetic. Thanks, however, to the masterful construction of her arguments, she succeeds not only in avoiding this but also in offering an engaging second look at "what is philosophy for" by confronting head-on the views she was arguing against. If only for this, the book is a must-read for apprentices of philosophic practice, for the jaded who presume either to have known enough philosophy or have known philosophy enough, and lastly, for those who do not think philosophy is worth knowing. Those who wish to reacquaint themselves with the question "What is philosophy for?" might also find this volume inviting. The simplest question is often the most elusive, to paraphrase Heidegger. Midgley's book shows us why and how this will remain a perennial predicament.

*Department of Philosophy, University of Santo Tomas, Philippines*

### References

Critchley, Simon, *Continental Philosophy: A Very Short Introduction* (Oxford: Oxford University Press, 2001).

Midgley, Mary, *What Is Philosophy For?* (London: Bloomsbury Academic, 2018).

---

<sup>7</sup> *Ibid.*, 200.

# Submissions

---

## Please Read Carefully

### A. What do we publish?

We are interested in publishing articles, review articles, and book reviews across the whole range of philosophical topics, but with special emphasis on the following subject strands:

- Filipino Philosophy
- Oriental Thought and East-West Comparative Philosophy
- Continental European Philosophy
- Anglo-American Philosophy

The journal primarily caters to works by professional philosophers and graduate students of philosophy, but welcomes contributions from other fields (literature, cultural studies, gender studies, political science, sociology, history, anthropology, economics, inter alia) with strong philosophical content.

We are not accepting creative works at the moment.

### B. How long should a submission be?

- Article (8,000 words or less)
- Review Article (8,000 words or less)
- Book Review (2,500 words or less)

### C. When should you submit and what happens to your submission??

We now accept submissions all year round. All unsolicited submissions undergo the following review process:

- Initial vetting process of the Editorial Board. In this stage, the Editorial Board screens based on their suitability for further review. All authors are informed about the result of this initial process. Please take note that a notice to proceed is not yet an offer to publish, it is merely to inform the author that his/her submission has passed the initial stage and will further undergo the blind peer-review process. It is also in this stage when the author is informed whether his/her submission is considered for the June or December issue.
- Blind peer-review process. When an author receives a notice to proceed, his/her submission will be sent to a nominated expert for blind review. In general, this process could take between two to six months, depending on various circumstances. Since this part of the review could be tedious, we solicit the patience of the author.
- Notice of acceptance, rejection, or provisional acceptance. Based on the referee's report, the author will be informed whether the Editorial Board has decided to accept or reject the submission for publication. In majority of cases, a provisional acceptance is given to an author who needs to revise his/her submission based on the recommendations of the referee and the Editorial Board. Full acceptance is only given when the recommended revisions are addressed.
- Notice to publish. A notice to publish will require the full cooperation of the author, as this stage involves editing the style, grammar, format, and overall layout of the accepted submission. Please take note that the Editorial Board reserves the right to exclude accepted submissions that do not comply with the stylistic standards of the journal.
- Release of the issue.

Because of the sheer number of unsolicited submissions that we receive on a daily basis, submission management has become a challenge for us. This often results in the piling-up of submissions, the breakdown of the online submission tool, and, at times, unacknowledged submissions. In this regard, we wish to solicit for the patience and full cooperation of contributors.

Recently, our online submission tool has been having an issue receiving submissions sent through Yahoo! Mail. We would like to request prospective contributors to refrain from using Yahoo! Mail when sending submissions through our online submission tool. Many thanks.

[Click here to go the Online Submission Tool.](#)

Please note that this form requires you to have a Gmail or G-suite account. If you do not have a Gmail or G-suite account, then you may submit directly to [submissions.kritike@gmail.com](mailto:submissions.kritike@gmail.com) with all the [required details](#). Thank you.



# Submissions

---

## Specific Submission Guidelines

1. We accept submissions in either English or Filipino with good punctuation, grammar, and spelling. Provide a 200-word abstract in English and at least 4 key words. Please take note of the number of the acceptable word count for your submission (see Section B above).
2. Kritike is a refereed journal and follows a double-blind review policy, which means that the identities of both the author and the referee are concealed during the review process. As such, please make sure that your manuscript is prepared for blind review, meaning your name and institutional affiliation should not appear in the body of your paper. If you cited your own previous work(s) in the article, delete your name from the citation(s).
3. We recommend that, at the first instance, you use our prescribed citation style. You may also use the Chicago style which resembles our own. **[Click here to visit the journal's style guide page.](#)**
4. Submit your text in 2.0 line spacing with 12 points font size. Quotations exceeding four lines should be indented and single-spaced.
5. Save your paper as either a Rich Text Format file (\*.rtf) or a Microsoft Word document (\*.doc or \*.docx).
6. We recommend that you submit your paper by filling in the **[online submission tool](#)** for a more systematic and efficient submission process.
7. By sending us your submission, you agree to be bound to the Terms and Conditions set in Section C of the journal's **Publication Ethics and Publication Malpractice Statement**.

# Publication Ethics and Malpractice Statement

---

KRITIKE is committed to meet the highest ethical standards in research and academic publication. The journal is guided by the following principles:

## A. Responsibilities of the Editorial Board

The Editorial Board ensures that manuscripts are prepared for blind peer-review. It is the responsibility of the Editorial Board to accept, reject, or recommend a manuscript for revision and resubmission. Such decision is based, to a large extent, on the recommendations of nominated experts who act as referees. It is the responsibility of the Editorial Board to inform an author about the status of his/her submission, regardless of the decision. The Editorial Board may choose to reject a paper that violates legal provisions on libel, copyrights, and originality (plagiarism). Information regarding a manuscript under review must remain confidential until it is finally accepted for publication. The Editorial Board does not necessarily endorse the views expressed in the articles published in the journal. **As an Open Access journal in the Gold category, KRITIKE does not charge any fees to complete the publication process. No charges are levied against the authors or users for submission or article processing.**

## B. Responsibilities of the Referee

The referees nominated by KRITIKE's Editorial Board are experts in their areas of specialization. The referees assist the Editorial Board's decision to accept, reject, or revise and resubmit manuscripts based on their objective assessments and recommendations. A referee must treat an assigned manuscript with utmost confidentiality during the peer review process; however, it is the responsibility of the referee to inform the Editorial Board when a legal violation by the author is suspected. The evaluation of a manuscript should be based solely on its academic merit and not on race, gender, sexuality, or religious and/or political orientation of the author.

## C. Responsibilities of the Author

It is the responsibility of the author to prepare his/her manuscript for blind review. The author must ensure that his/her work is original and not plagiarized. The sources used in the manuscript should be properly cited. An author must not submit the same manuscript to another journal when it is currently under review by KRITIKE. It is the responsibility of an author to inform the Editorial Board right away if his/her manuscript is being considered in another journal or publication medium; in such case, KRITIKE will discontinue the review of the manuscript. If an author's manuscript is published by KRITIKE, he/she must adhere to the provisions set in the **Copyrights** section of the journal

## Contact Us

---

If you wish to send us your feedback, general questions about the journal, questions about article submissions, theme suggestions for future issues, or word of intention to be a peer-reviewer or referee, send a message to [kritike.editor@gmail.com](mailto:kritike.editor@gmail.com).

If you wish to be a peer-reviewer or referee, do send us your complete name, e-mail address, institutional affiliation, position, and area of expertise via e-mail (include subject heading: reviewer). If you have any suggestions for specific themes (e.g., “European Philosophy and the Filipino Mind” or “Is there such thing as Filipino Philosophy?”) for future issues of the journal, send them via (include subject heading: theme).

Please note that unsolicited submissions should be sent through the journal's Article Submission Tool.

You can also contact us via snail mail:

KRITIKE

c/o **Dr. Paolo A. Bolaños**  
Department of Philosophy  
Room 109 Main Building  
University of Santo Tomas  
España, Manila 1015  
Philippines