Volume 9 Issue 2 2022

Symposion

Theoretical and Applied Inquiries in Philosophy and Social Sciences

Romanian Academy, Iasi Branch
„Gheorghe Zane" Institute for Economic and Social Research
ISSN: 1584-174X EISSN: 2392-6260

Contact address: Institutul de Cercetări Economice și Sociale „Gh. Zane", Iași, str. T. Codrescu, nr.2, cod 700481, Romania. Phone/Fax: 004 0332 408922. Email: symposion.journal@yahoo.com. www.symposion.acadiasi.ro.

TABLE OF CONTENTS

RESEARCH ARTICLES

RESEARCH ARTICLES

Extending the Is-ought Problem to Top-down Artificial Moral Agents

Robert James M. Boyles

Abstract: This paper further cashes out the notion that particular types of intelligent systems are susceptible to the is-ought problem, which espouses the thesis that no evaluative conclusions may be inferred from factual premises alone. Specifically, it focuses on top-down artificial moral agents, providing ancillary support to the view that these kinds of artifacts are not capable of producing genuine moral judgements. Such is the case given that machines built via the classical programming approach are always composed of two parts, namely: a world model and utility function. In principle, any attempt to bridge the gap between these two would fail, since their reconciliation necessitates for the derivation of evaluative claims from factual premises.

Keywords: artificial moral agent, David Hume, is-ought problem, machine ethics, top-down AMA.

Introduction

In *Hume's Law as another Philosophical Problem for Autonomous Weapons Systems*, Boyles (2021) argues that military-grade autonomous weaponry are susceptible to the is-ought problem. Autonomous Weapons Systems (AWS), on the one hand, may be defined as machines that, "once activated, can select and engage targets without further intervention by a human operator." (Department of Defense 2012, 13) The is-ought problem, on the other hand, is a logical problem commonly attributed to David Hume (Gunkel 2018, 88). Also known as "Hume's Law" (Hare 1952) or "Hume's Guillotine" (Black 1964, 166), the said problem espouses the thesis that evaluative conclusions may never be inferred from factual premises alone (Restall and Russell 2010, 243).

For Boyles, grounding the purported moral judgements of AWS appear to be intractable in light of the is-ought problem, since these artifacts make use of factual data from their environments to carry out specific actions. Supposedly, the process involved in such requires the derivation of evaluative statements from a set of purely factual ones. He further explains that, "[i]f there remains a fundamental difference between the actions or behaviors of … AWS from their human counterparts – the latter being capable of arriving at genuine moral judgments, then one should be cautious in naively trusting the apparent ethical decisions of the former." (2021, 126)

Though certain distinctions between human beings and AWS were pointed out in the said article, particularly in terms of ethical decision-making and their moral standing, no supplementary explanations were offered to account for why

the latter are unable to arrive at genuine ethical decisions (i.e., beyond just appealing to Hume's no-ought-from-is doctrine). For one, the notion that there are different ways of designing intelligent machines was not considered. Thus, this paper looks into how the is-ought problem relates to the internal design of artifacts, specifically focusing on top-down artificial moral agents.

An artificial moral agent (AMA) is commonly defined as "an artificial autonomous agent that has moral value, rights and/or responsibilities." (Sullins 2009, 208) In order to realize this type of machine, the top-down method of designing AMAs subscribes to the view that moral principles may be directly encoded into its internal program (Wallach and Allen 2009, 83-97). By doing so, an artifact's actions and behaviors would, thus, be regulated by the said precepts. However, note that several challenges have also been put forward against the top-down AMA track.

Apart from the difficulty of translating and applying abstract moral principles to specific, actual situations, Misselhorn (2018, 165) holds that top-down AMAs are predisposed to the frame problem. In a nutshell, the latter problem concerns logic-based systems, specifically on how to represent the effects of their actions. Supposedly, identifying the particular conditions in modeled environments that have been affected by the actions of top-down systems pose certain hurdles, since there is an assumption that all other conditions stay fixed. The said assumption, however, is still unfounded, citing the unresolved issue of aptly sorting out all relevant information from the irrelevant ones. Dennett (1984, 130), for one, asserts that this issue eventually results in a "deep epistemological problem."

Allen, Smit, and Wallach (2005), on the other hand, explain that a major concern with top-down AMAs is that the ethical rules or commandments programmed into them often conflict with one another, especially once the said systems encounter real-world ethical dilemmas. They further maintain that such conflicts result in computationally intractable situations, and most all rule-based systems do not offer concrete ways to resolve them.

The primary aim of this paper is to raise another challenge against top-down AMAs, which is the is-ought problem. Following Boyles' (2021) use of the latter to proffer certain foundational worries against AWS, the present work extends the said strategy to top-down AMAs. In this paper, it is argued that the said systems are also prone to Hume's Law, since machines built via the said method are always composed of two parts, namely: a world model and utility function (Hall 2011, 512). In principle, any attempt to bridge these two parts would fail, since reconciling them would be the same as deriving evaluative statements from a set of factual ones. Furthermore, note that, although Hall (2011, 514) briefly mentions that the world model and utility function of classical systems are separated by 'Hume's is-ought guillotine,' no extensive explanation for such has been provided. Hence, this article also seeks to offer a more detailed analysis of the said idea.

To contend that top-down AMAs are susceptible to the no-ought-from-is doctrine, the following section initially revisits Hume's original discussion of the said problem, while also citing the two views that resulted from it (i.e., moral descriptivism and moral non-descriptivism). The main objective of this part is to highlight the idea that there seems to be no foolproof solution to the is-ought problem today. The subsequent section, meanwhile, provides a summary on the view that AWS are susceptible to the is-ought problem. Moreover, in order to further ground this notion, specifically on how it relates to top-down AMAs, the next section looks into the nature of classically programmed artifacts. In this section, it is shown that the reason why top-down technologies are unable to produce genuine moral judgements is that the world model and utility function embedded in them, in principle, cannot really be reconciled. The final section of this paper provides a few concluding remarks.

Hume's No-ought-from-is Doctrine

As mentioned earlier, the origins of the no-ought-from-is doctrine may be traced to Hume (Gunkel 2018). In his *Treatise on Human Nature*, Hume states that:

> In every system of morality, which I have hitherto met with, I have always remark'd, that the author proceeds for some time in the ordinary way of reasoning, and establishes the being of a God, or makes observations concerning human affairs; when of a sudden I am surpriz'd to find, that instead of the usual copulations of propositions, is, and is not, I meet with no proposition that is not connected with an ought, or an ought not. This change is imperceptible; but is, however, of the last consequence. For as this ought, or ought not, expresses some new relation or affirmation, 'tis necessary that it shou'd be observ'd and explain'd; and at the same time that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it. But as authors do not commonly use this precaution, I shall presume to recommend it to the readers; and am persuaded, that this small attention wou'd subvert all the vulgar systems of morality, and let us see, that the distinction of vice and virtue is not founded merely on the relations of objects, nor is perceiv'd by reason. (1739/1964, 243-244)

It could be inferred from the above quote that the is-ought problem centers on the viability of providing factual justifications for moral judgments. For Hume, no legitimate deduction[1] may be made from an 'is' to an 'ought.' (Brown 2008, 229)

Following Hume's line of thinking, it may be said that in any argument that is composed of (1) a set of purely factual premises and (2) a normative conclusion (i.e., derived from the said set), the normative judgment found in the latter would not logically follow from the factual assertions found in the series of is-statements.[2] Snare (1992, 84-85) also explains that the is-ought problem may be

[1] Hume's usage of the term 'deduction' has resulted in a debate on what he truly meant by this. See Schurz (1997, 2). The present work is neutral about this issue.

[2] The is-ought problem may also be further related to Hume's view on ethics. See Cohon (2010).

likened to an in-principle thesis – on the foundational level, evaluative conclusions may never be arrived at as long as one deduces them from factual premises alone. To further understand this, one may cite the logical relationship between the nature of 'oughts' and 'isses.'

One way to account for is-statements is to think of them as the content of assertions or descriptive statements. Note that the latter are truth-evaluable expressions, and a standard example of such are declarative sentences. [3] Conversely, ought-statements operate more like imperatives or prescriptions of actions.[4] So, in contrast to is-statements, ought-statements cannot be evaluated as either true or false, since they do not state facts.[5]

Since ought-statements generally pertain to a moral obligation or a norm of conduct (i.e., in the context of moral judgments), they naturally relate to the notion of ethical value. This is because all moral systems normally presuppose a close link between moral obligations and ethical values (Schurz 1997, 1). The idea behind this is that what is deemed as ethically good ought to be, must be, or needs to be done, which demonstrates the obligatory aspect, if not the imperative force, of an ought claim. So, ought-statements function more as prescriptions of actions, and they contrast with is-statements that bear truth claims.

Ever since Hume pointed out the apparent logical invalidity of deducing 'oughts' from 'isses,' a number of philosophers have proffered different ways to address the no-ought-from-is doctrine. Among the numerous replies to the latter include that of Hare (1952) and Searle (1964), which could be treated as standard representatives of the universal prescriptivist and moral descriptivist views, respectively.

Prescribing Ought-statements

As discussed by Boyles (2021, 118), universal prescriptivism adheres to the notion that ought claims are a kind of prescription or imperative (Gensler 2011, 56). In *The Language of Morals*, Hare contends that the "language of morals is one sort of prescriptive language." (1952, 1) So, for prescriptivists, imperatives do not really state facts, which further means that these can neither be true nor false.

Prescriptions operate like commands, basically directing someone to do or perform something. For prescriptivists, ought-statements are just expressions of impartial desires of how one should live, act, or behave (Gensler 2011, 57). In a sense, this demonstrates the normative aspect of imperatives, which also accounts for why prescriptivists believe that prescriptions are universalizable.

---

[3] As also noted by Boyles (2021, 126), not all declarative sentences are is-statements. Furthermore, not all kinds of assertions can be judged as straightforwardly true or false.

[4] Though there is an ongoing debate as to whether or not ought-statements are, in fact, truth-evaluable, this issue is well beyond the scope of this work.

[5] Several philosophers, like Hume, have argued that moral judgments do not, strictly speaking, state facts, which makes them non-truth-evaluable also. For a brief summary of Hume's view regarding this issue, see Cohon (2010).

Ought-statements, for prescriptivists, are universalizable prescriptions (Gensler 2011, 58). If one comes up with an 'ought,' then this does not merely equate to the act of making a simple prescription. Putting forward an ought-statement expresses one's utmost desire that an action be context-invariant (i.e., the prescribed course of action ought to be followed in all analogous cases). The said idea is also embodied in the logical rules for 'oughts', which are as follows:

> U. To be logically consistent, we must make similar evaluations about similar cases.
>
> P. To be logically consistent, we must keep our moral beliefs in harmony with how we live and want others to live. (Gensler 2011, 58)

Logical rules U and P are consistency rules for action (Gensler 2011, 58). Logical rule U dictates the iteration of a specific action in all analogous cases. This means that whenever an ought-statement is made, we should treat its content as context-invariant. On the other hand, logical rule P maintains that ought judgments are, in fact, imperatives, which further entails that an ought judgment is somehow devoid of its obligatory function (i.e., in the moral sense). Gensler further notes that, "[i]nstead, they tell us what we must do if we're to be logically consistent in our moral beliefs." (2011, 58) This highlights the notion that an ought-statement becomes a logical test for the consistency of our moral judgments and beliefs.

As for the issue of universal prescriptivism being a rational ethical system, even though it regards ought-statements as non-truth bearing claims, Gensler (2011, 57) maintains that it is quite possible to construct a system comprised of a set of prescriptions that does not necessarily equate to a moral system. He further asserts that among the said systems include cookbooks, the laws of a particular country, and complex computer programs, to name a few.

Gensler (2011, 61-63), however, holds that prescriptivism may further lead to the denial of the possibility of attaining moral knowledge and truths. If ought-statements are just universalizable prescriptions, then moral judgments would only be a test of consistency of prescribed actions. In a way, this highlights the idea that no further moral truths may be attained given that, for there to be further moral truths, new information must be accounted for.

In relation to the is-ought problem, prescriptivists readily affirm such. As discussed earlier, Hare was even famous for coining the phrase "Hume's Law." (Cohon 2010) Considering that they accept Hume's Law, the only recourse for prescriptivists is to show that, in all moral arguments, there is an underlying evaluative statement hidden, if not assumed, alongside the relevant factual premises (Joaquin 2012, 55-56). So, with regard to the attempt of deriving an ought-conclusion from a series of is-statements, it appears that prescriptivism does not yield a tenable solution to the is-ought problem at present. For prescriptivists, the said problem is, in fact, a live one.

Moral Talk as Factual Claims

Proponents of moral descriptivism argue that ethical language is best treated as an attempt to describe something in the world (Fisher and Kirchin 2006, 3). As per Boyles (2021, 118-120), ethical statements are, for them, simply reducible to claims about facts (i.e., under a certain set of conditions). Thus, moral statements could also be evaluated, like descriptive statements, based on their truth-value. To further grasp the descriptivist model, specifically in the context of how it deals with the no-ought-from-is thesis, consider Searle's (1964) view regarding this matter.[6]

To address the is-ought problem, Searle first challenges the notion that facts are entirely distinct from values (Joaquin 2012, 56-57). He demonstrates this by providing the following counterexample:

1) Jones uttered the words 'I hereby promise to pay you, Smith, five dollars.'
2) Jones promised to pay Smith five dollars.
3) Jones placed himself under (undertook) an obligation to pay Smith five dollars.
4) Jones is under an obligation to pay Smith five dollars.
5) Jones ought to pay Smith five dollars. (Searle 1964, 44)

As to how one may derive the evaluative claim, "Jones ought to pay Smith five dollars," from the said set of factual statements, Searle (1964, 44-49) explains that this could simply be done by adding 'empirical assumptions, tautologies, and descriptions of word usage' to the given premises (1964, 48).

Moreover, by using definitional connections between 'promise,' 'obligate,' and 'ought,' as well as including a *ceteris paribus* clause to eliminate possible contrary considerations, Searle claims that the move from premises (2) to (5) seems "relatively easy." (1964, 49) Recall that moral descriptivists, like Searle, hold that ought-statements could be reduced into fact-stating propositions under a given set of conditions. For Searle, he is able to specify such conditions by employing the concept of institutional facts (Joaquin 2012, 65).

In a nutshell, institutional facts are specific kinds of facts that depend on human convention and agreement (Searle 1995, 29). In contrast to brute facts (e.g., Water is H_20 in this world), which exist independently of human agreement (Searle 1995, 27), institutional facts presuppose human institutions, since they are responsible for creating the system of constitutive rules – those that not only regulate, but also ensure the rules' very existence.

Searle employs the idea of institutional facts to specify the scope or conditions that allows for the translation of ought-statements into descriptive

[6] It should be pointed out here that Searle was actually responding to the more modern formulation of the is-ought problem, which was put forward by philosophers such as Hare (Joaquin 2012, 56).

ones. So, for instance, it may be argued that the statement "Jones ought to pay Smith five dollars" may be considered true only if one presumes that there exists a human or social institution that states that such is the case. As ingenious as Searle's strategy may seem, however, a couple of concerns may be raised against it (Boyles 2021, 119).

First, Searle's maneuver to infer an 'ought' from an 'is' is not without problems. In fact, he tries to anticipate many objections to this.[7] For example, his supplementary premise to bridge (4) and (5), "(4a) Other things are equal," may be rendered to the following statement: "(4a) All those who are under an obligation, *ceteris paribus*, ought to fulfill that obligation." Note that this seemingly equal formulation could be treated as an (implicit) ought premise, which would put into question Searle's primary goal of deducing an 'ought' from purely is-statements.[8]

Second, even if one grants that Searle is successful in deriving an 'ought' from a set of 'isses,' it must be pointed out that such strategy appears to only work for a very particular case, specifically, to promise making (Boyles 2021, 119). Given the seemingly limited scope of the said strategy, it may be argued that it is really not that fruitful in light of the endeavor of developing autonomous machines.

With regard to the attempt of deriving an 'ought' from a series of is-statements, it appears that universal prescriptivism and moral descriptivism do not yield, as of yet, a tenable solution to this issue. As mentioned earlier, both strategies are not ironclad. Furthermore, note that the worries against these two positions have also been related to the prospect of creating autonomous weapons systems.

Hume's Guillotine, Autonomous Weaponry and Moral Judgments

As regards the development and deployment of AWS, many have already called for more research into the ethical concerns and dangers surrounding these types of technologies (Sharkey 2010; Sparrow 2016; Boyles, Dacela, Evangelista, and Rodriguez 2022, 192). Boyles (2021), for one, proffers that such are prone to the no-ought-from-is doctrine.

To establish that AWS are susceptible to Hume's is-ought, Boyles (2021, 115) first cites Boulanin and Verbruggen's (2017, 7-11) idea that the concept of autonomy in artifacts is basically operationalized by integrating three fundamental capabilities (i.e., sense, decide, and act). He further explains that the sense capacity, mainly composed of built-in sensors and sensing software, is utilized by AWS to perceive the environment (Boyles 2021, 120). So, all data generated by this capacity are particular facts about the context an AWS is presently situated, and these facts become input for the decide capability. After a

[7] See Searle 1964, 49-52.

[8] This worry against Searle's strategy may be generally classified to fall under the "objections against the *ceteris paribus* clause." (Joaquin 2012, 59-63)

specific decision has been reached, an autonomous system, then, implements a set of actions. So, in the sense-decide-act cycle, data input is critical, since the judgements and actions of AWS depend on the information gathered by its sensors and sensing software.

As per Boyles, the actions of AWS that subscribe to the sense-decide-act cycle cannot be trusted, morally speaking, since there is no direct way of reconciling their sense capacity with the decide part (2021, 120). This is because doing so requires inferring evaluative statements from a set of factual ones, which entails that all of their supposed moral judgments are of no ethical worth. He further maintains that attempts to get around the said issue, like encoding AWS with ethical precepts, would not work as such strategy is similar to the prescriptivist solution to the is-ought problem.

Recall that, for universal prescriptivists, adding an evaluative statement to a series of descriptive ones would circumvent the is-ought problem. This is because prescriptivists claim that this additional statement would enable the deduction of an evaluative conclusion from the said total set of premises. So, in the context of AWS, a machine would be capable of generating a moral judgment, *prima facie*, as long as it is pre-programmed with certain ethical principles. However, Boyles (2021, 121) forewarns that this might be deceiving.

Citing Gensler (2011, 57), Boyles explains that it is possible to develop a model consisting of a set of prescriptions, but is devoid of any ethical value. Standard examples of these kinds of systems include the laws of different nations, computer programs, and the like. So, with regard to prescriptivist-based AWS, it might be the case that their apparent ethical judgments do not really have any moral worth.

Furthermore, in terms of developing descriptivist-based AWS, Boyles claims that this strategy is problematic as well. He states that:

> Programming the decide capacity of an AWS so that it could decide which particular theory is the most relevant in a specific situation somehow issues in the frame problem. As per the said problem, tagging a theory as the most relevant one necessitates for an artifact to consider infinitely finite facts in a given situation... Note also that the task of ascertaining which ethical theory is most appropriate in a given context parallels the descriptivist solution to Hume's Law. This is because descriptivists claim that moral statements are reducible to factual ones under specific conditions, perhaps arguing that a moral theory may be considered most relevant in a given situation when it addresses the issue at hand. (2021, 124)

Since descriptivist solutions to the is-ought problem still make use of implicit ought-claims (i.e., to generate evaluative conclusions from descriptive statements), Boyles (2021, 124) notes that proposals of this type also do not stand on firm grounds. Additionally, recall that descriptivist solutions, like that of Searle (1964), appear to only work in very limited circumstances. If the objective is to create robust and autonomous machines, then making use of the descriptivist design strategy appears to be futile.

Considering the issues of both prescriptivist and descriptivist solutions to the is-ought problem (i.e., in relation to the endeavor of developing AWS), Boyles (2021) claims that there is a fundamental difference between machines and their human counterparts, especially in terms of their moral standing and ethical decision-making. He maintains that, at present, we are certain that humans could enact genuine moral judgments, but it is not that clear if AWS, or any other kinds of machines, could actually do the same (2021, 123).

In light of the said standing issues against prescriptivism and descriptivism, Boyles (2021) holds that grounding the apparent moral judgments of AWS is problematic. If the ultimate goal of creating these types of machines is for them to become autonomous agents, then there is no certainty that they would always come up with ethically-sound actions (i.e., in the context of real-life moral dilemmas). This is because the actions of AWS largely depend on the data gathered by their sensors and sensing software, and the process of generating the former from the latter parallels the move of deriving evaluative conclusions from a set of purely factual premises.

Note that Boyles (2021), however, does not really provide any further explanation as to why AWS are not capable of producing genuine ethical judgments (i.e., beyond just appealing to Hume's is-ought). For one, the notion that there are different types of artificial moral agents, which are intelligent artifacts that have the capacity to enact moral decisions (Cervantes et al. 2020), was not considered.

For Wallach and Allen, with regard to developing AMAs, three design strategies could be employed, namely: the top-down, bottom-up, and hybrid approaches. As explained earlier, the top-down AMA method adheres to the idea that moral principles may be directly encoded into an artifact's internal program (2009, 83-97). By programming certain ethical precepts from the get-go, the actions and behaviors of these kinds of machines would be regulated, morally speaking.

Bottom-up AMA approaches, on the other hand, are those that employ evolutionary, learning, or developmental methodologies (Wallach and Allen 2009, 80). This track focuses on creating environments where artifacts could consider and enact different courses of action, while also learning from them in the process. Note that bottom-up AMAs are given set rewards whenever they exhibit praiseworthy behavior. In contrast, hybrid AMAs integrate the design principles of both top-down and bottom-up options, and one way to supposedly realize this method is by using Aristotelian virtue ethics (Wallach and Allen 2009, 10). Wallach and Allen note that "[v]irtues are a hybrid between top-down and bottom-up approaches, in that the virtues themselves can be explicitly described, but their acquisition as character traits seems essentially to be a bottom-up process." (2009, 10)

Considering the different AMA design tracks, it might be the case that some further explanation is needed as to why AWS, if not all AMAs, are unable to

generate genuine ethical decisions. After all, with regard to providing moral capacities to artifacts, Misselhorn highlights such differences, explaining that:

> [An] important issue is how moral capacities can be implemented in artificial systems. This entails two questions: first, with which moral standards artificial systems should be furnished and, second, how those standards can be implemented. Both issues are related since a decision for a certain ethical framework also entails certain constraints on its realization in a software program. (2018, 165)

Furthermore, Misselhorn notes that, as regards artificial systems, there are three approaches to moral implementation (i.e., the top-down, bottom-up, and hybrid tracks), and these differing methods "bring together a certain ethical theory with a certain approach to software design." (2018, 165) So, perhaps the particular ways that the no-ought-from-is doctrine is explicitly and individually realized in the different AMA tracks should also be considered. For the purposes of the present work, the top-down AMA approach is further examined in relation to the is-ought problem.

Top-down AMAs and Classical Programming

As noted earlier, the top-down approach for designing AMAs could be realized through encoding ethical principles into a machine's internal program. The assumption here is that, once certain moral precepts have been hardwired into the latter, the actions and behaviors of artifacts would, then, be regulated by such. Wallach and Allen further explain that:

> ... a top-down approach takes an ethical theory, say, utilitarianism, analyzes the informational and procedural requirements necessary to implement this theory in a computer system, and applies that analysis to the design of subsystems and the way they relate to each other in order to implement the theory. (2009, 80)

So, for Wallach and Allen (2009, 84), the top-down track centers on the idea of having a set of rules that, in turn, could be developed into an algorithm. Note that the foundational assumptions of this track, in a way, may also be related to the direct programming or classical AI method of creating intelligent systems.

Proponents of classical AI, also known as 'Symbolic AI,' assert that artificially intelligent systems may be achieved by writing sophisticated computer programs (Haugeland 1985, 112-114). [9] This parallels Searle's idea that an "appropriately programmed computer really is a mind, in the sense that computers given the right programs can be literally said to understand and have other cognitive states." (1980, 417) Note that the said view, which Searle calls

[9] Haugeland (1985, 112) also calls the classical AI track as 'Good Old Fashioned Artificial Intelligence' or GOFAI.

'strong AI,' states that the human mind could be likened to a computer program, implemented by a brain that functions as its hardware (Mabaquiao 2014).[10]

Classical AI works under the assumption that brains are nothing more than complex machines, which entails that, in order to create autonomous agents[11] (Pfeifer and Scheier 1999, 25-27), one has to write computer programs that would serve as their central intelligence system. If one, then, grants that the foundational assumptions of the top-down AMA track are, in principle, the same with, if not grounded on, classical AI, the next issue would be how to show that the former is susceptible to Hume's is-ought problem. To address this issue, one may look into the inner workings of a top-down AI's artificial mind.

To argue that AMAs developed via the top-down method are indeed predisposed to the is-ought problem, it must be recalled that these technologies have built-in computer programs that try to mimic the human mind, specifically its ability to exercise thinking. In addition, it should be highlighted that such programs are based on the notion that an artificial mind is always composed of two parts, namely: (1) a world model and (2) utility function (Hall 2011, 512).

A world model (WM), on the one hand, is the part of an artificial mind that houses the objective knowledge regarding the world. It contains all the facts about the world that it is modeling, and this may, in turn, be used by an artifact to plan, evaluate, and predict the effects of its actions. The utility function (UF), on the other hand, "establishes a preference between world states with which to rank goals." (Hall 2011, 515) It may, thus, be said that the WM of an artifact is the one primarily responsible for providing it with the current state of affairs in the world, including the different possibilities or consequences of its actions, while its UF calculates which of these is the most desirable given a specified goal.

The problem with the said model is that there seems to be no way of bridging the gap between a machine's utility function and its world model. This is because, in principle, the move from first modeling the actual and potential state of affairs in the world, to finally deciding which among these possibilities is most preferable, is quite similar to the attempt of deriving an 'ought' from a series of factual statements about the world.[12] The is-ought problem is at work in this model, and there seems to be no clear solution to this philosophical worry at present, as discussed in the previous sections.

Recall that the WM is composed of various facts about the world. Suppose that an autonomous artifact with a central processing system, for instance, sees a speeding automobile that appears to be in a collision course with a group of tourists. Its WM would generate a number of facts about the situation, including

[10] This is in contrast to what Searle has called "weak AI," (1980, 417) which holds that computers are nothing but powerful tools for studying the mind.

[11] This may be further related to what Goertzel has dubbed as "artificial general intelligence." (2007, 1161-1163)

[12] As noted earlier, Hall (2011, 514) also mentions that the WM and UF are separated by 'Hume's is-ought guillotine,' but there appears to be no detailed analysis of the said idea.

the potential casualty count if nobody does anything, the risks involved in helping, and so on. From all of these facts, the artifact may come up with predictions of the possible effects and side effects of its action, if not its inaction.

However, note that the procedure by which a machine decides that a particular action is better over another seems to be quite untenable. This is because, citing the no-ought-from-is thesis, evaluative conclusions may never be derived from factual statements, which is said to be the manifest function of the UF in making sense of the factual data generated by a machine's WM.

It may then be argued that, if AMAs developed via the top-down route subscribe to the WM-UF model, then these artificially intelligent machines, strictly speaking, would not really be able to come up with genuine moral judgments due to the is-ought problem. Either their so-called judgements are absent of good moral grounding, or their conclusions are somehow empty in the ethical sense. Note that this further explains Boyles' (2021) point that technologies based on the sense-decide-act cycle are quite questionable in terms of their moral standing and ethical decision-making.

On top of the idea that AWS are unable to produce genuine moral judgments because their actions largely depend on the data gathered by their sensors and sensing software, it may be argued that all technologies modeled via the direct programming method would not be able to perform the said task given their internal design. These artifacts are unable to generate actions with actual moral worth because there is no way of reconciling their WM-UF parts. Such is the case given that this step would require the deduction of evaluative conclusions from purely factual premises.

If one takes into consideration the internal design of top-down AMAs, the idea that certain kinds of intelligent machines are, in principle, predisposed to the is-ought problem would make further sense. Consider, for instance, Boyles' (2021, 124-125) explanation why descriptivist-based AWS are not capable of circumventing Hume's Law. As regards the said notion, Boyles states that descriptivist strategies still make use of implicit ought-claims (i.e., to come up with an evaluative conclusion from descriptive premises). Furthermore, he also holds that, "if such is the case, then it may also be argued that the descriptivist solution to Hume's Law is nothing different from the prescriptivist idea that evaluative judgments may be uncovered in the factual premises of moral arguments." (2021, 124) In a way, this explanation regarding the said types of AWS become more intelligible if one further relates such to the fact that top-down AMAs are largely determined by their world model and utility function.

Consider the following case: suppose that a top-down AMA finds itself in a position of having to confront a modified version of Foot's (1967) trolley problem. Imagine a runaway train that is fast speeding down a railway and there are five individuals at the end of one of the tracks. The said train is headed right straight for these people, and the top-down machine could prevent their demise because it is standing by the lever that controls the tracks. If it pulls the lever, the train

would switch to a different set of tracks. However, if the said lever is pulled, the train would head directly to another person on this different track. What must our top-down AMA do? Should it pull the lever? If it does, the train would be diverted onto the new track where one person would be killed. If it does not, the train would kill the five individuals at the end of the main track. Which is the correct choice?[13]

How would a top-down AMA address the mentioned situation? Let us suppose that our artifact has built-in ethical systems (e.g., utilitarianism, deontology, etc.) from which it could choose the right course of action to take. But the said artifact has a problem: "What ethical theory should it choose to base the right course of action?" To answer this, let us suppose that, on top of its basic program, there is a second-order platform that might guide our artifact to favor or prefer one ethical theory over the rest. Here, an impasse is reached.

All considerations in the second-order platform seem to be exactly the same as those in the first. Since there is no forthright way of deciding which moral principle is better over the other regardless of the programming levels (i.e., without, of course, begging the question), this endeavor would likely lead to a problem of circularity. Boyles (2021, 123) further explains that the said strategy might even be prone to relativism, if not an arbitrary assignment of values – preferring a specific theory, but with no justified (moral) grounds. Note also that the issue of having a multitude of ethical theories that conflict with each other has already been raised previously (Tonkens 2009; Lara and Deckers 2019).

By stressing the difficulties in refereeing between opposing moral precepts, which in effect also highlights the issue of employing ought-premises in programming ethical machines, it could be said that one is left to work with only factual propositions. Note that such a case eventually results in yet another is-ought problem. However, it might still be argued by others that the adjudication process between the various competing ethical theories may be addressed by simply giving a machine a certain modification. For one, Boyles' (2021, 124) considers the possibility of this issue being "addressed by further adjusting the decide capacity" of an artifact. Actually, taking into account the WM and UF of top-down AMAs, the said modification concerns an artifact's UF, while its WM might also be affected.

Instead of just focusing on a machine's decide capacity, citing a top-down AMAs internal program would provide a better picture on why they are predisposed to Hume's is-ought. In altering a machine's UF, note that labeling a particular theory as the most appropriate one (i.e., as compared to other ethical precepts) denotes that such a theory is actually the most relevant among its competitors. So, claiming that the decide capacity of an artifact would be the one

[13] Note that the original intention of Foot's (1967) thought experiment is to show that there is a difference between letting someone die and killing a person. This, for her, has ramifications on the moral status of some abortion cases.

affected by the proposed strategy (Boyles 2021, 124) appears to have only scratched the surface.

Furthermore, with regard to Boyles' (2021, 124) view that the process of identifying something as being the most relevant would eventually run into issues, note that a top-down AMA's MW could also be examined to better understand this. For one, remember that he generally maintains that encoding the decide capacity of an AWS (i.e., so that it could adjudicate and select which particular theory is most relevant in a given situation) somehow leads to the frame problem. It may, thus, be said that this issue concerns a machine's WM, since the latter would be the one responsible for modeling each and every fact about a specific situation, resulting in an infinite regress. This is the reason why an artifact with such a program would not be able to generate genuine moral judgments; this task entails that a top-down machine would have to infinitely account for all the factual data processed inside its WM.

Recall that descriptivists claim that moral statements may be reduced to factual premises under certain conditions (Boyles 2021, 118). So, it may be argued that an ethical theory is the most relevant one if it is the most apt in a given situation. However, it must be remembered that the descriptivist approach is doomed to fail, since it still smuggles in implicit ought claims (i.e., in attempting to infer an 'ought' from a series of is-statements) as part of its starting set of facts.

Moreover, even if one grants that this strategy succeeds, it seems to only work in very narrow cases at best. It may even be argued that the possibility of actually identifying such narrow cases is close to impossible because this exercise could lead up to other issues, like the frame problem – considered by many as a technical and philosophical problem that focuses on "representing the effects of action[s] in logic without having to represent explicitly a large number of intuitively obvious non-effects." (Shanahan 2016) Note that this parallels the view of Moss, that "[d]etermining the best action at every moment would overwhelm a finite computational device." (2016, 2) Yet again, it seems that further looking into the internal design of top-down AMAs (i.e., that they are largely determined by their WM and UF) provides further grounding as to why certain types of machines are unable to produce genuine ethical decisions.

Finally, a couple of things may also be pointed out about the idea of pre-programming ethical theories into AMAs. First, it must be noted that, even if the top-down track prospers, it would not be that simple to assign concepts like 'praiseworthiness,' 'blameworthiness,' and so on to such artifacts. Since the said values were just pre-programmed to them by their designers, achieving full moral agency by means of this track is a bit questionable. For one, Krzanowski and Trombik explain that:

> Can then such a deep ethics be computed (in the Church–Turing sense), given that metaphysics is not mathematical? Ethical rule-based on Hobbesian, Kantian, utilitarian or other ethical schools can be to some extent translated into a computer algorithm and made 'computable.' But then all 'metaphysical'

> dimensions of the ethical actor are 'lost in translation.' If a machine is programmed according to 'translated' rules... this ethics would be a special type of ethics, not ethics in the deep, metaphysical sense. (2021, 143)

Krzanowski and Trombik hold that ethics, in a deep (metaphysical) sense, is non-computable, and they maintain that there really is no other way of defining and accounting for what is "computable." (2021, 143)

Recall that, although one may construct a system consisting of a set of prescriptions, like a standard computer program, this system does not really equate to a moral system (Gensler 2011, 57). So, whenever AMAs built through the top-down route initiate actions that, at first glance, appear to be ethical in nature, such as their apparent moral judgments, it might be the case that the actions of these machines are actually devoid of any ethical value.

Second, note that there is a difference between ethical reasoners and ethical decision-makers (McDermott 2008). Ethical reasoners are artifacts that model the reasoning processes of human beings (i.e., in coming up with ethical conclusions). Ethical decision-makers, in contrast, are those that duplicate or mimic what in people are classified as ethical decisions. The primary distinction between these two is that the latter really understands what is actually at stake whenever moral dilemmas or conflicts arise (e.g., the ethical thing to do in a given situation and how it seems to conflict with one's self-interest).

On the other hand, ethical reasoners, in a way, just mechanically generate moral conclusions by, say, considering the facts at hand. Similarly, Hunyadi explains that:

> As far as *machine ethics* is concerned, this means one thing: if you program a specific set of ethical principles into a machine, you do not make the machine an artificial *moral* agent, but an *executor of those specific principles*, which is an entirely different thing. This so-called 'artificial agent' will be expected to respond according to *those ethical principles*, chosen by the programmer. (2019, 62)

Hunyadi, thus, further clarifies that, as regards an artificial system, it is more apt to label such an 'artificial moral executor' instead of 'artificial moral agent,' specifically an "artificial *utilitarian*, *deontological* or *perfectionist* executor, depending on the ethical principles chosen by the programmer." (2019, 62)

In light of the differences cited above, it could be contended that, in order to actually realize the concept of moral machines, they should not only be simple ethical reasoners, but ethical decision-makers as well. Unfortunately, the suggested strategy of encoding machines with pre-programmed ethical precepts does not fall under the latter. So, the idea of assigning ethical notions such as praiseworthiness, blameworthiness, and others to artificial moral agents built via the top-down track seems like a lost cause.

Conclusion

As regards top-down AMAs, it was argued that they would not be able to come up with morally-relevant judgments, since artifacts built this way are primarily composed of two parts, namely: a world model and utility function. In principle, there is no way of reconciling these two parts because such would entail the derivation of evaluative claims from a set of purely factual ones, and this goes against the general tenets of the is-ought problem. One consequence of this is that either the so-called moral judgements of these AMAs would be absent of any good moral grounding, or their generated conclusions would be empty of ethical value. Note that, instead of simply citing Hume's is-ought, as well as the sense-decide-act cycle, one could further look into a machine's internal design in order to have a better understanding as to why they are incapable of moral reasoning.

As discussed above, there are different types of artificial moral agents, developed either via the top-down, bottom-up, and hybrid approaches. With regard to top-down AMAs, which is the focus of the present work, it appears that they may not be considered as authentic moral agents (i.e., as compared to humans) due to their internal design. Note, however, that there are those, like Nadeau (2006), who also contend that even biological humans may not be considered as moral agents.

Prospectively, further research on how the other forms of AMA technologies, specifically those designed via the bottom-up and hybrid methods, fare against the no-ought-from-is thesis may be looked into. Would these strategies be susceptible to Hume's is-ought as well? If yes, then how would this relate to other concerns put forward against such types of AMAs? For instance, in relation to bottom-up AMAs, consider Baum's view that "it is impossible for AI designers to avoid embedding certain ethics views into an AI... because there is no one single aggregate ethical view of society." (Baum 2020, 167) It is, thus, interesting to know if the in-principle thesis embedded in the is-ought problem could aid in further understanding these sorts of ideas. Of course, if it really turns out that all these AMAs encounter insurmountable issues resulting from Hume's Guillotine, then perhaps it may be high time to look for other viable alternatives.

References

Allen, Colin, Iva Smit, and Wendell Wallach. 2005. "Artificial Morality: Top-down, Bottom-up, and Hybrid Approaches." *Ethics and Information Technology* 7: 149-155.

Baum, Seth D. 2020. "Social Choice Ethics in Artificial Intelligence." *AI & Society* 35 (1): 165-176.

Black, Max. 1964. "The Gap Between 'Is' and 'Should.'" *Philosophical Review* 73 (2): 165-181.

Boulanin, Vincent, and Maaike Verbruggen. 2017. *Mapping the Development of Autonomy in Weapon Systems*. Solna: Stockholm International Peace Research Institute.

Boyles, Robert James M. 2021. "Hume's Law as another Philosophical Problem for Autonomous Weapons Systems." *Journal of Military Ethics* 20 (2): 113-128.

Boyles, Robert James M., Dacela, Mark Anthony, Evangelista, Tyrone Renzo, and Jon Carlos Rodriguez. 2022. "COVID-19 and Singularity: Can the Philippines Survive another Existential Threat?" *Asia-Pacific Social Science Review*, *22* (2), 181-195.

Brown, Charlotte R. 2008. "Hume on Moral Rationalism, Sentimentalism, and Sympathy." In *A Blackwell Companion to Hume*, edited by Elizabeth S. Radcliffe, 219-239. Oxford: Blackwell Publishing Ltd.

Cervantes, José-Antonio, Sonia López, Luis-Felipe Rodríguez, Salvador Cervantes, Francisco Cervantes, and Félix Ramos. 2020. "Artificial Moral Agents: A Survey of the Current Status." *Science and Engineering Ethics* 26: 501-532.

Cohon, Rachel. 2010. "Hume's Moral Philosophy." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Available at: http://plato.stanford.edu/entries/hume-moral. Accessed: 11 August 2021.

Dennett, Daniel. 1984. "Cognitive Wheels: The Frame Problem of AI." *In Minds, Machines, and Evolution: Philosophical Studies*, edited by Christopher Hookway, 129-151. Cambridge: Cambridge University Press.

Department of Defense. 2012. *Autonomy in Weapon Systems*. DOD Directive 3000.09. Washington: Department of Defense. Available at: https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf. Accessed: 3 August 2021.

Fisher, Andrew, and Simon Kirchin. 2006. *Arguing about Metaethics*. Oxon: Routledge.

Foot, Philippa. 1967. "The Problem of Abortion and the Doctrine of Double Effect." *Oxford Review* 5: 5-15.

Gensler, Harry J. 2011. *Ethics: A Contemporary Introduction.* 2nd ed. New York: Routledge.

Goertzel, Ben. 2007. "Human-level Artificial General Intelligence and the Possibility of a Technological Singularity: A Reaction to Ray Kurzweil's The Singularity is Near, and McDermott's Critique of Kurzweil." *Artificial Intelligence* 171: 1161-1173.

Gunkel, David J. 2018. "The Other Question: Can and Should Robots have Rights?" *Ethics and Information Technology* 20: 87-99.

Hall, Storrs J. 2011. "Ethics for Self-Improving Machines." In *Machine Ethics*, edited by Michael Anderson and Susan Leigh Anderson, 512-523. Cambridge: Cambridge University Press.

Hare, Richard Mervyn. 1952. *The Language of Morals.* Oxford: Clarendon Press.

Haugeland, John. 1985. *Artificial Intelligence: The Very Idea*. Cambridge: MIT Press.

Hume, David. 1739/1964. *A Treatise of Human Nature*, edited by L.A. Selby-Bigge. Oxford: Clarendon Press.

Hunyadi, Mark. 2019. "Artificial Moral Agents. Really?" In *Wording Robotics: Discourses and Representations on Robotics*, vol. 130, edited by Jean-Paul Laumond, Emmanuelle Danblon and Céline Pieters, 59-69. Cham: Springer.

Joaquin, Jeremiah Joven. 2012. "John Searle and the Is-ought Problem." *Scientia* 2 (1): 53-66.

Krzanowski, Roman M., and Kamil Trombik. 2021. "Ethical Machine Safety Test." In *Transhumanism: The Proper Guide to a Posthuman Condition or a Dangerous Idea?*, edited by Wolfgang Hofkirchner and Hans-Jörg Kreowski, 141-154. Cham: Springer.

Lara, Francisco, and Jan Deckers. 2019. "Artificial Intelligence as a Socratic Assistant for Moral Enhancement." *Neuroethics*, 1-13. (please check reference, as it is inexact: https://philpapers.org/rec/LARAIA-5)

Mabaquiao, Napoleon. 2014. "Turing and Computationalism." *Philosophia: International Journal of Philosophy* 15 (1): 50-62.

McDermott, Drew. 2008. "Why Ethics is a High Hurdle for AI." In *North American Conference on Computers and Philosophy*. Bloomington, Indiana, July 9-12. Available at: http://www.cs.yale.edu/homes/dvm/papers/ethical-machine.pdf. Accessed: September 16, 2014.

Misselhorn, Catrin. 2018. "Artificial Morality. Concepts, Issues and Challenges." *Social Science and Public Policy* 55: 161-169.

Moss, Henry. 2016. "Genes, Affect, and Reason: Why Autonomous Robot Intelligence Will Be Nothing Like Human Intelligence." *Techné: Research in Philosophy and Technology* 20 (1): 1-15.

Nadeau, Joseph Emile. 2006. "Only Androids Can Be Ethical." In *Thinking about Android Epistemology*, edited by Kenneth M. Ford, Clark Glymour and Patrick Hayes, 241-248. Cambridge: MIT Press.

Pfeifer, Rolf, and Christian Scheier. 1999. *Understanding Intelligence.* Cambridge: MIT Press.

Restall, Greg, and Gillian Russell. 2010. "Barriers to Consequence." In *Hume on Is and Ought*, edited by Charles Pigden, 243-259. Basingstoke: Palgrave Macmillan.

Schurz, Gerhard. 1997. *The Is-ought Problem: An Investigation in Philosophical Logic.* Dordrecht: Springer.

Shanahan, Murray. 2016. "The Frame Problem." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Available at: https://plato.stanford.edu/entries/frame-problem/. Accessed: 24 October 2017.

Sharkey, Noel. 2010. "Saying 'No!' to Lethal Autonomous Targeting." *Journal of Military Ethics* 9 (4): 369-383.

Searle, John R. 1964. "How to Derive 'Ought' From 'Is.'" *The Philosophical Review* 73 (1): 43-58.

————. 1980. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3 (3): 417-457.

————. 1995. *The Construction of Social Reality*. London: Penguin Books Ltd.

Snare, Francis. 1992. *The Nature of Moral Thinking.* London: Routledge.

Sparrow, Robert. 2016. "Robots and Respect: Assessing the Case Against Autonomous Weapon Systems." *Ethics & International Affairs* 30 (1): 93-116.

Sullins, John P. 2009. "Artificial Moral Agency in Technoethics." In *Handbook of Research on Technoethics*, edited by Rocci Luppicini and Rebecca Adell, 205-221. New York: IGI Global.

Tonkens, Ryan. 2009. "A Challenge for Machine Ethics." *Minds and Machines* 19 (3): 421-438.

Wallach, Wendell, and Colin Allen. 2009. *Moral Machines: Teaching Robots Right From Wrong.* New York: Oxford University Press.

The Private Language Argument: Another Footnote to Plato?

Arnold Cusmariu

Abstract: A valid and arguably sound private language argument is built using premises based on Wittgenstein's *Philosophical Investigations* augmented by familiar analytic distinctions and concepts of logic. The private language problem and the solution presented here can be plausibly traced to Plato's Allegory of the Cave. Both literatures missed the connection.

Keywords: Plato, Wittgenstein, epistemic privacy, following a rule, outward criteria, type-token distinction, substitution theorem.

I. Coming Full Circle

The story how this article came to be written may be of interest to some readers. Those who do not wish to travel down memory lane with me can skip ahead.

In *Critique of Pure Reason*, Kant famously wrote:

> Human reason has this peculiar fate that in one species of its knowledge it is burdened by questions which, as prescribed by the very nature of reason itself, it is not able to ignore, but which, as transcending all its powers, it is also not able to answer.

The question that 'burdened' me since learning geometry in middle school that I was 'not able to ignore' but that 'transcended my powers to answer' was a practical one: How do mathematical proofs work? I had to do them in school and, like other students, sometimes I got it right and sometimes I didn't. So, what did I do well in one case but not the other? My teachers did not explain. They would only tell us when we got it right and when not. Logic was not taught as a subject in its own right.[1] In fact, I didn't even know there was such a subject until very much later. When I asked a mathematics professor in college to explain the concept of mathematical proof, he answered that the question was not mathematical but philosophical. I changed majors.

On to Philosophy!

Socratic Method and Critical Philosophy by a neo-Kantian, Leonard Nelson – a book I happened across in a Dover Publications catalog – confirmed that I had made the

[1] I'm hoping that my book *Logic for Kids: All Aboard the* ***Therefore*** *Train*, Jenny Stanford Publishing (forthcoming) will change that situation. This is the book I wish I'd had when I was ten.

right decision. Philosophy asked fascinating questions, some of which seemed familiar for some reason. However, it was the challenges and allure of Ludwig Wittgenstein's oracular *Tractatus Logicus-Philosophicus* that got me hooked.

- Proposition 1.1 informed the reader that "the world is the totality of facts, not of things." Well and good but a totality is a set, which raises questions (a) how membership is to be determined and (b) whether reality was an abstraction of some sort because sets are abstract. But then, maybe I was wrong to understand 'totality' in mathematical terms.
- Proposition 1.1 also seemed to imply that Genesis ontology was wrong: that in the beginning, God created not things like heaven and Earth as the Bible asserted but rather facts! Here was a philosopher, and an obscure one at that (at the time), challenging a basic tenet of the most famous and influential book in the world, read and taken literally by millions for centuries. In olden times, such daring would have been considered a capital offense and punished accordingly.
- I thought Proposition 2 should have read "What is the case – a fact – is the *occurrence* of states of affairs," distinguishing *existence* from *occurrence*. After all, apples and oranges *exist* (ghosts and unicorns don't) but events and states of affairs *occur* (or don't). Such sentences as "the apple is red" and "the apple is red and not red" are both meaningful, hence states of affairs correspond to them but only one can occur.
- Are there facts as well as states of affairs according to Proposition 2; or is Proposition 2 a definition of 'fact,' reducing facts to states of affairs? Hmmm…
- I thought Proposition 2.0123, "If I know an object, I also know all its possible occurrences in states of affairs" entailed skepticism because we can't know all possible occurrences of an object in states of affairs.
- What does 'objects occur in a states of affairs' mean, anyway?
- No doubt I was missing something. I had to keep reading to find out.

It was helpful that my teachers at the City College of New York, Michael Levin, Arthur Collins, Charles Evans, and Daniel Bronstein were sympathetic to Wittgenstein. I eventually read *Philosophical Investigations*, did well in my philosophy courses and was admitted to the Honors program in my senior year, which required writing a thesis. I produced "Verification as a Family Resemblance Concept," a Wittgensteinian critique of positivist meaning in A.J. Ayer's *Language, Truth and Logic*. Levin supervised.

Armed with a *cum laude* degree with honors in philosophy, I had high hopes as I applied to Ivy League universities – Harvard, Princeton, Yale and Brown – to pursue graduate study. I also applied to the University of Oregon because Frank Ebersole, a Wittgensteinian Collins admired, taught there. Brown and Oregon accepted me. I decided on Brown, which proved to be a great decision indeed. It shaped my intellectual development in a variety of ways I did not appreciate at the time and also kept me out of the horrific mess known as the Vietnam War.

Providence must have persuaded the draft board official to sign off on my exemption. I was off to Providence, Rhode Island.

Philosophy at Brown University

I knew Roderick Chisholm taught at Brown but had not studied his books[2] and was unaware of his view that we have privileged access to our mental states in an epistemic sense, a view he shared with Descartes. When I raised Wittgenstein's well-known point at *Investigations* 246 that we can't be said to know that we are in pain, a student in the Chisholm seminar replied, dismissively: "That's just something Wittgenstein said."[3] Yes, Wittgenstein did say that but 'just' implied that he had not given reasons for this view, which was false. I then wrote a paper for Chisholm explaining what I understood to be Wittgenstein's reasons. I wish I'd kept it! I did keep a paper I wrote for him titled "Are There Propositions?" and I'm glad I did. He made various comments in red pencil, including an extremely generous one for which I will be forever grateful: "Very good indeed – just the way a philosophical paper ought to be written." I took the advice to heart and have tried to follow it ever since.

I studied epistemology with Chisholm; philosophy of science with Sosa; Kant with Van Cleve; Hume with Lenz; logic with Luschei; philosophy of language with Swartz; and ethics with Brock. No seminars were offered on Wittgenstein, so the private language argument never came up.[4] Vincent Tomas taught aesthetics, published on the subject (Tomas 1959), and had an interest in Wittgenstein. I wish I'd taken his seminar. I often wondered how (or whether) acquaintance with philosophical problems about art would have shaped my development as an artist years later. As it turned out, however, metaphysics and epistemology had a lot more to do with it.

In matters of method, Chisholm agreed with Plato that definitions were necessary if only to avoid pointless verbal disputes by making sure everyone was on the same page. Technical concepts were routine in science and mathematics, why not philosophy? This approach was a welcome antidote to the linguistic analysis promoted by Wittgenstein's *Philosophical Investigations*, which regarded philosophical problems as something to be dissolved rather than solved, sort of like untying a knot – showing a fly a way out of the fly bottle, as Wittgenstein put it in passage 309.

Chisholm also agreed with Plato in matters of ontology. He accepted predicable universals such as properties and relations, drawing a sharp distinction between *existence* and *exemplification*. He also accepted non-predicable universals such as events and states of affairs, drawing a sharp

[2] *Perceiving: A Philosophical Study* was published in 1957 and the first edition of *Theory of Knowledge* in 1966. The directly evident takes up Chapter 2 of *Theory of Knowledge*.
[3] A Brown student before my time had a different take on this issue. See Canfield 1975.
[4] Chisholm commented on the private language argument briefly in Hahn 1977, 27.

distinction between *existing* and *occurring*. Brown helped me put the family resemblance model in the rear-view mirror, where it has stayed since.

Dissertation Time

I worked on the problem of universals for my dissertation, one of the oldest and most fundamental in philosophy, flirting initially with an Aristotelian solution – criticized later in Cusmariu 1979B – ultimately proposing a Platonist solution. The dissertation showed that an ancient problem, Plato's 'Third Man' Argument (TMA), and a modern one, Russell's Paradox (RP) – the property version of which I learned about from Chisholm – were opposite sides of the same coin: one was about self-predication, the other about non-self-predication. Accordingly, the same solution should handle both. How exactly that could be done was a lot of hard work, though intellectually very exciting.

I realized right away that Wittgensteinian approaches to TMA and RP were not an option. There was no way to 'dissolve' such complex and fundamental problems by examining 'what we would say' in the vernacular about self-predication and non-self-predication. These were technical concepts that seldom found their way in common parlance and as such provided scant usage data to analyze.[5]

Ah, if only Plato had pondered if there was a form exemplified by all and only forms that did not self-exemplify! On the other hand, Russell, who knew the history of philosophy very well indeed and was a Platonist at one time, nevertheless did not see the connection between RP and TMA. Had he seen it, he might have wondered if his theory of types also solved the TMA.[6] His theory of descriptions wouldn't have.

Lasting Impact

Platonism has been with me ever since. It had significant impact much later when I began making sculpture.[7] I realized that a radically different approach to art, practically as well as theoretically, could be gleaned from Plato's beautiful metaphor of *interweaving forms* in *Sophist*. Several articles explain my working aesthetic – Cusmariu 2009, 2015A and B, and 2017A and B – which could not have been written without the training I got at Brown. I take seriously Plato's objections to art in Book X of *Republic* and strive to respond to them in my sculptures. I addressed these objections in Cusmariu 2016A.

---

[5] See Cusmariu 1978, 1979A, and 1985 for expositions of key results of my dissertation.

[6] Wittgenstein's *Tractatus* also failed to see the connection between TMA and RP. His knowledge of the history of philosophy was reportedly shaky, so it's unlikely he was even aware of TMA. At *Tractatus* 3.333, Wittgenstein claims to have 'disposed' of the property (propositional function) version of RP. It is debatable whether his proposal 'disposes' of the set paradox. It does nothing to resolve the TMA.

[7] The last section has a photo of a recent sculpture of mine.

So, here I am full circle, back to Wittgenstein, having followed in the meantime methodological guidelines set by Plato, with an assist from Chisholm.

II. Platonist Guidelines

We may infer from the sign that tradition says was displayed above the entrance to Plato's Academy, "Let None but Geometers Enter Here," that applicants were expected to possess significant reasoning skills acquired from study of mathematics, which they would need to apply to study philosophy. Four key elements were involved:

- **Defining a problem**

In geometry, this entailed formulating theorems about properties of points and lines, collected by Euclid in his *Elements*.

- **Motivating the problem**

The importance of theorems taken singly is determined by the contribution they make to enlarging the system's storehouse of knowledge.

- **Solving the problem**

Solving the problem means presenting a proof, which Euclid did for every theorem he listed.

- **Defending the solution**

Defending the proof entailed soundness and validity. Soundness is assumed to come from self-evident axioms. Validity is taken for granted because mathematics considers the logic of proofs to be self-evident (ha!) and as such not in need elaboration. This has been true in mathematics since Euclid, whose *Elements* does not include what we now call proof theory, e.g., rules of inference and replacement and a substitution theorem. Logic has been a silent partner in mathematical pedagogy ever since (see Cusmariu 2016B.)

At Plato's Academy:

- **Defining the problem** meant asserting a significant thesis – e.g., that there was more to knowledge than true belief; that there was more to justice than self-interest; that art was a dangerous illusion; and so on – then formulating necessary and sufficient conditions for key concepts such as knowledge and justice that went beyond the merely stipulative definitions of geometry.
- As to **motivating philosophical problems**, the problem of universals, which Plato was the first to formulate, is one of the most fundamental problems in philosophy, cutting across virtually all fields of our discipline. His solution, the Theory of Forms, has profound implications.[8]
- To spell out a **problem-solving methodology**, Plato offered the hypothesis-refutation model and *reductio ad absurdum* form implicit in the Socratic Method, both as part of dialectic (see below) – the philosopher's tool box.

[8] Succinct statements of the problem of universals and Plato's solution may be found in Cusmariu 1979.

- As to **reasoned defense of solutions**, had Aristotle been a geometer – how did he get into the Academy? – he would have realized that his 15 valid categorical syllogisms in standard form, effectively rules of inference, are inadequate for the purpose of making dialectic logically explicit. Plato himself could do no better, alas. Some scholars believe he tried and failed.

Argumentation Impact

There are four counterparts in philosophical argumentation:

> **I**: Formulating a proposition as the conclusion of an argument.
>
> **II**: Motivating the philosophical significance of that proposition.
>
> **III**: Listing premises claimed to logically support the conclusion.
>
> **IV**: Showing that the argument is valid and that premises withstand scrutiny.

The TMA, which took dead aim at Plato's own Theory of Forms, satisfied the first three requirements remarkably well all things considered but requirement IV was a bridge too far that early in the history of logic, still too far even after Aristotle. Nevertheless, TMA defined an important dialectical requirement – taking a hard critical look at one's own theory – that philosophy has followed ever since. TMA also represents a technical innovation far ahead of its time, for which Plato has not been given credit, namely, that components of a system must form consistent set.[9] We know now that proving consistency poses serious problems that, unfortunately, Euclid and mathematicians after him never even considered – until forced to do so by Frege, Russell and Whitehead.

III. Cutting to the Chase

Having set the stage with some admittedly unusual twists and turns, we are ready for passages associated with the private language argument (PLA) in Wittgenstein's *Philosophical Investigations* (PI). According to the literature, passages from 243 to 380 of PI contain material from which an argument can be assembled for the proposition that a private language is logically impossible.[10] Restricted for the time being to sensations, the PLA can be cast as an argument of the following form:

> W1. Theory T is true only if it is logically possible to refer to sensations by means of a private language.

[9] TMA components that do not form a consistent set are: (1) there is a unique form of *F-ness* exemplified by all and only *F* objects; (2) all and only *F* objects are in the same class; (3) *F-ness* itself is an *F* object, and (4) *F-ness* itself is not in the class of all and only *F* objects. A resolution of the inconsistency is in Cusmariu 1980. Gilbert Ryle claimed in his 1939 article that exemplification leads to an infinite regress. This problem is solved in Cusmariu 1985.

[10] Notable exceptions to this view are Baker 1998, Stroud 2000 and Canfield 2001.

W2. It is logically possible to refer to sensations by means of a private language only if condition C can be satisfied.

W3. Condition C cannot be satisfied.

Therefore, by *modus tollens*,

W4. It is not logically possible to refer to sensations by means of a private language.

Therefore, by *modus tollens*,

W5. Theory T is false.

The PLA literature, which is vast and growing, has sought to explain:

(a) what a private language is,

(b) what privacy in general means,

(c) what referring to sensations means,

(d) what condition C might be,

(e) whether condition C is necessary as claimed in W1,

(f) whether condition C can be satisfied as contested in W2,

(g) which answers to (a)-(f) are consistent with PI.[11]

A thorough discussion of these matters is a book-length project. The best that I can do in this paper is to put together an argument using a minimum of assumptions based on the PI text, without claiming, however, that my PLA reconstruction necessarily represents Wittgenstein's intent. A brief Q&A section at the end addresses some problems with my minimalist reconstruction.

IV. Meeting Argumentation Requirements

Reconstruction of a PLA is the best that I can do because the PI text does not meet the argumentation requirements enunciated earlier.

Requirement I: Conclusion?

The proposition that a private language is logically impossible, which is considered to be the conclusion of a PLA, is not stated in PI passages that have been associated with the argument. That is, there is no German sentence in these pasages that reads "eine private Sprache ist logisch unmöglich." While the phrase 'private language' (German, *private Sprache*) occurs several times, it does not occur as part of a sentence that expresses the proposition that a private language is logically impossible. Also absent is language that states, implies or even hints

[11] See Candish and Wrisley 2019.

that such a proposition is to be the conclusion of an argument in the standard sense.

Requirement II: Motivation?

A useful way of becoming familiar with a scholarly work is to check the names of people and titles of books listed in the index.

- The only philosophers cited in PI are Augustine, Gottlob Frege, William James, Frank Ramsey, Bertrand Russell and Socrates (but not Plato). G.E. Moore is cited in the context of his paradox.
- The only books cited in PI are Plato's *Theaetetus* (without authorship attribution) and Wittgenstein's own *Tractatus*.
- The literature[12] considers Descartes and British Empiricists – Locke, Berkeley and Hume – as principal targets of a PLA, yet none of these names or their works are listed in the PI index.

Requirement III: Premises?

There are no premises identified as such in passages the literature has associated with a PLA; nor is there even a hint as to what assumptions are to belong together as a valid and sound defense of the proposition that a private language is logically impossible. Another standard component of arguments, a claim that premises logically imply the conclusion, is also not expressed. This problem affects the entire book. While arguments appear to be proposed in various passages, none are spelled out and would require significant work to translate into standard form.

Requirement IV: Validity?

Because standard argumentation components are absent, it is no surprise that there is no formal defense of PLA in PI. Well, if the object is to 'dissolve' philosophical problems, why would there be?

IV: Kripke Demurs

Saul Kripke, author of a controversial book on Wittgenstein (Kripke 1982) begs off satisfying argumentation requirements that are standard in analytic philosophy, walking on eggshells as he explains his approach to PI (Kripke 1982, 5):

> I suspect – for reasons that will become clearer later – that to attempt to present Wittgenstein's argument precisely is to some extent to falsify it. Probably many of my formulations and recastings of the argument are done in a way Wittgenstein would not himself approve.

[12] See, for example, Jones 1971, editorial introduction, 13-15.

Approve? I doubt that Kripke would assert that precise presentations of arguments by other philosophers (Aristotle? Hume? Kant?) risked 'falsifying' them. Kripke does not explain why he thinks Wittgenstein deserves special treatment, though this attitude is nothing new (see biographical articles in Fann 1967, 11-130).

C.D. Broad, a professor at Cambridge in the 1930s, says he skipped weekly Moral Sciences Club meetings in part because (Schilpp 1959, 61):

> I was not prepared to spend hours every week in a thick atmosphere of cigarette-smoke, while Wittgenstein punctually went through his hoops, and the faithful as punctually 'wondered with a foolish face of praise.'

Gilbert Ryle, teaching at Oxford, noticed the same phenomenon when he attended Moral Sciences Club meetings (Monk 1990, 495):

> Veneration for Wittgenstein was so incontinent that mentions, for example, my mentions, of any other philosopher were greeted with jeers.

We do not, and indeed should not, worry whether even major figures in the history of philosophy would have 'approved' of this or that interpretation of their work. Though Wittgenstein tells us in the Preface of PI that "my thoughts soon grew feeble if I tried to force them along a single track against their natural inclination," that's not a reason we should take a philosopher's personal predilections (excuses?) seriously. We should focus on the details of his contribution and asses rigorously whether it has philosophical merit. I view philosophy as a collection of problems in an eternal, Platonist sense, to which solutions are offered from time to time – some good, some bad, some indifferent – so that the personal preferences are only of biographical significance, if that. Let's find out what solutions are being proposed and get on with it.

V. PLA Reconstruction Overview

My reconstruction is based on enough PI passages to make attribution reasonable. It relies also on the type-token distinction due to Peirce (Peirce 1906) and the Substitution Theorem (ST) due to Frege (see LeBlanc 1966):[13]

- The literature has identified a total of 137 passages, beginning at 380 and ending at 243, on which a PLA reconstruction can be based, to which I will add three more, 201, 202 and 580, for a potential total of 140 passages.
- It turns out, however, that far fewer passages are needed to build an argument for the proposition that a private language is logically impossible. Accordingly, the resulting argument can be termed 'minimalist.'
- The type-token distinction (TTD) is a familiar one in analytic philosophy and requires no explanation. I will assume without argument that TTD can apply to

[13] Such a theorem is also necessary to show that a substitution instance of a valid syllogism is itself valid, which Aristotle never recognized. No one else did until Frege.

linguistic as well as non-linguistic objects without worrying about the difference between various applications. I will bracket the issue whether there is version of TTD that is consistent with Wittgenstein's family resemblance model.

- An informal statement of ST will be sufficient for present purposes:
- ST. Uniform substitution of propositional letters in a valid argument form ((P1, P2, P3 … Pn) / C) will result in a valid argument.

Here is a pictorial explanation of uniform substitution for *modus ponens* from my book *Logic for Kids*:

(…) → (…)
(…)

(…)

VI. Private Tasks

In the ordinary sense of 'private performance,' a private performance is given before an exclusive audience, sometimes an audience of one; for example, a performance of a piano sonata by Beethoven before a head of state. A violinist practicing for a recital with no one listening can also be considered a private performance in an extended ordinary sense even though there is no audience.

A private performance in the sense relevant to a PLA is also a performance in which the performer and the audience are one and the same. However, such a performance is exclusive in another sense. Privacy as applied to performing a task in the sense relevant to a PLA is exclusive in an epistemic sense that is conceptually similar to privacy as applied to language in an epistemic sense. Thus, performing a task privately for purposes of a PLA means performing it according to a method whose application only the performer can understand. A key objective of PLA is to challenge the possibility of carrying out certain tasks privately in this sense.

Note, however, that it is the application of the method that is private in the requisite sense, not the method itself, which need not necessarily be private in any sense. So, here is a schema for expressions of the form "person P can perform task X privately," to be instantiated and expanded as we proceed.

PT. Person P can perform task X privately =df Only P can understand how he himself is able to apply method M to perform task X.

- Violin playing can be private in the ordinary sense of being heard by a limited audience (or no audience) but not in the sense of PT. Left hand fingering technique, for example, is not such that only the performing violinist can understand how he is able to apply this technique to playing the violin.
- The method M most compatible with PI is probably introspection. A PLA need not challenge the meaningfulness of introspection as a concept nor its

application in general; only its application to carrying out certain tasks, identified below.

VII. Passage 202

Passage 202 will figure prominently in my PLA reconstruction, contrary to the literature (except Kripke 1982). However, I find existing translations problematic. I need to take care of this matter before proceeding any further.

Here is the German original followed by translations by G.E.M Anscombe in the third edition of PI (202A)[14]; by P.M.S. Hacker and Joachim Schulte in the fourth edition of PI (202B)[15]; and my own translation (202C).

> **202**: Darum ist 'der Regel folgen' eine Praxis. Und der Regel zu folgen *glauben* ist nicht: der Regen folgen. Und darum kann man nicht der Regel 'privatim' folgen, weil sonst der Regel zu folgen glauben dasselbe wäre, wie der Regel folgen.
>
> **202A**: And hence also 'obeying a rule' is a practice. And to *think* one is obeying a rule is not to obey a rule. Hence it is not possible to obey a rule 'privately'; otherwise thinking one was obeying a rule would be the same thing as obeying it.
>
> **202B**: That's why 'following a rule' is a practice. And to *think* one is following a rule is not to follow a rule. And that's why it's not possible to follow a rule 'privately'; otherwise thinking one was following a rule would be the same thing as following a rule.
>
> **202C**: That's why 'following a rule' is a practice. And to *believe* one is following a rule is not to follow a rule. And that's why one cannot follow a rule 'privately'; otherwise believing one was following a rule would be the same as following it.

Translation Comments

- The meaning of "folgen" in English is "to follow." The German of "to obey" is "gehorchen," which does not occur in the German original of 202.
- The meaning of "glauben" in English is "to believe." The German of "to think" is "denken," which does not occur in the German original of 202.
- The PI text contains dozens of occurrences of "denken" and its cognates, all of which translators render literally as "to think" and its cognates. I found no translations of "denken" as "to believe."
- Occurrences of "glauben" are translated literally as "believe" at 24, 105, 140 and 260. In addition to 202, occurrences of "glauben" translated as "think" can be found at 69, 101, 114 and 139.
- The meaning of "kann man nicht" in English is "one cannot." The German of "it is not possible" is "es ist nicht möglich," which does not occur in the German original of 202.
- Therefore, 202C is a more accurate translation.

[14] Wittgenstein, Ludwig. 1958 [1953]. *Philosophical Investigations*, 3rd edition, translated by G.E.M. Anscombe. London: Blackwell.

[15] Wittgenstein, Ludwig. 2009 [1953]. *Philosophical Investigations*, 4th edition, translated by G.E.M Anscombe, P.M.S. Hacker and Joachim Schulte. Revised by P.M.S. Hacker and Joachim Schulte. London: Blackwell.

- Belief ascriptions rely on the dispositional sense of 'belief,' whereas 'thought' is closer to an occurrent sense of mental activity. Wittgenstein appears to intend the dispositional sense, which is another reason 'belief' should be used in translating 202.
- Beliefs are bearers of truth values along with statements and sentences, not thoughts. For example, the traditional definition of knowledge is justified true belief, not justified true thought.
- The modal operator raises the question whether the possibility in the passage is logical or metaphysical, which is best avoided in the PI context.
- More seriously, because the 'hence' after the second sentence suggests an inference, deriving "it is impossible to bring about X" from "one cannot do X" attributes an obvious howler to Wittgenstein!

VIII. Passages Relevant to My PLA Reconstruction

This section lists PI passages relevant to my version of PLA, along with comments as appropriate.

> **201**: That's why there is an inclination to say: every action according to a rule is an interpretation. But one should speak of interpretation only when one expression of a rule is substituted for another.

Substantive Comment: This remark hints at the Substitution Theorem. Wittgenstein was familiar with Frege's work and may have known ST. My PLA reconstruction does a good deal more with this important theorem than is hinted at in 201.

> **202** (my translation): That's why 'following a rule' is a practice. And to *believe* one is following a rule is not to follow a rule. And that's why one cannot follow a rule 'privately'; otherwise believing one was following a rule would be the same as following it.

Substantive Comment 1: The text of 202 seems to suggest the following valid argument: (1) If person P can follow a rule privately, then P thinking that he is following a rule logically implies that he is following a rule. (2) Thinking that-p does not logically imply that-p. Therefore, (3) P cannot follow a rule privately. The issue is why the first premise is true, which cannot be explained merely by appealing to whatever definition of privacy is appropriate for following a rule privately.

Substantive Comment 2: Premise (2) is not obvious if construed about seemings. "It seems to me that I am following rule R" does not logically imply "I am following rule R"; but "it seems to me that I am having toothache" does logically imply "I am having a toothache." I found no passage in PI that explains the difference.

Substantive Comment 3: Following a rule privately is problematic according to my PLA reconstruction for a very different reason than the one the text indicates.

> **243**: The words of this language are to refer to what only the speaker can know – to his immediate private sensations (German, *auf seine unmittelbaren, privaten, Empfindungen*.) So another person cannot understand the language (German, *diese Sprache*).

Substantive Comment 1: 'The language' in the second sentence has a more restricted scope than it might appear. Only words that perform a denotative function are at issue, namely, words that refer to a person's 'immediate private sensations.'

Substantive Comment 2: There is an ambiguity in 243 that can be cleared up using the type-toke distinction. An 'immediate private sensation' is a sensation token. Pain felt at a specific time in a specific part of the body for a specific duration is a sensation token.

> **244**: But how is the connection between the name and the thing named set up? This question is the same as: how does a human being learn the meaning of the names of sensations? – of the word 'pain' for example. Here is one possibility: words are connected with the primitive, the natural, expressions of the sensation and used in their place.

Substantive Comment: The context of Plato's Cave Allegory in Book VII of *Republic* links up with passage 244 for reasons explained below.

> **256**: Now, what about the language which describes my inner experiences and which only I myself can understand? How do I use words to signify (German original, *bezeichne*) my sensations? – As we ordinarily do? Then are my words for sensations tied up with my natural expressions of sensations? In that case my language is not a 'private' one. Someone else might understand it as well as I. – But suppose I didn't have any natural expression of sensation, but only had sensations? And now I simply *associate* names with sensations, and use these names in descriptions.

Translation Comment 1: A literal translation of the German original "*Wie* bezeichne ich meine Empfindungen mit Worten?" is "How do I refer to my sensations with words?"

Substantive Comment 1: Translators give Wittgenstein (undeserved) credit for a key distinction not drawn in the text between language in an abstract sense and making use of a specific type of linguistic expression for a specific purpose. A minimalist PLA reconstruction need not take a stand on the issue of privacy of language in an abstract sense; only on the issue of making private use of linguistic expressions that can perform a denoting function. We need to tread much more carefully here than the PI text indicates.

Substantive Comment 2: Type-token ambiguity is also present in passage 256. I do not know if Wittgenstein was familiar with Peirce's distinction. The family resemblance model suggests that he wasn't or didn't think it was philosophically significant. I haven't researched what the literature says about this, if anything.

> **257**: So, does he understand the name, without being able to explain its meaning to anyone? – But what does it mean to say that he has 'named his pain'? – How has he managed this naming of pain? And whatever he did, was it on purpose? – When one says "He gave a name to his sensation", one forgets that much must be prepared in the language for mere naming to make sense.

Substantive Comment: The context of Plato's Cave Allegory links up with passage 257 as well, for reasons explained below.

> **258**: Let us imagine the following case. I want to keep a diary about the recurrence of a certain sensation.

Substantive Comment 1: Recurrence is a key component of the problem of universals (see Cusmariu 1979B), the others being predication and classification. Interestingly, PI does not invoke the family resemblance model to explain sensation recurrence.

Substantive Comment 2: Recurrence makes sense only for sensation types. Sensation tokens do not recur.

> To this end I associate it with the sign 'S' and write this sign in a calendar for every day on which I have the sensation.

Substantive Comment 3: The 'it' must be a sensation token, so that the passage should be read as ending with 'on which I have the sensation token.'

> – I first want to observe that a definition of the sign cannot be formulated. – But all the same, I can give myself a kind of ostensive definition. – How? Can I point to the sensation? – Not in the ordinary sense. But I speak, or write the sign down, and at the same time I concentrate my attention on the sensation – and so, as it were, point to it inwardly. – But what is this ceremony for? For that is all it seems to be! A definition surely serves to establish the meaning of a sign, doesn't it? – Well, that is done precisely by the concentrating of my attention; for in this way I impress on myself the connection between the sign and the sensation. – But "I impress it on myself" can only mean: this process brings it about that I remember the connection *correctly* in the future. But in the present case I have no criterion of correctness. One would like to say: whatever is going to seem correct to me is correct. And that only means that here we can't talk about 'correct.'

Substantive Comment 4: Descartes would have made the correct point that if it seems to to person X that he is experiencing a certain sensation-token, then it is at least evident to X that he is experiencing that sensation-token and perhaps even true that he is experiencing that sensation token, so X does not need a criterion of correctness. To be perceived is to be for sensation tokens. Descartes point remains even if the Evil Genius is manipulating a person's brain to produce a phantom sensation token. A toothache does not feel any less awful for being simulated by the Evil Genius. Phantom pains are no less painful because the leg is missing.

> **259**: Are the rules of the private language *impressions* of rules? – The balance on which impressions are weighed is not the *impression* of a balance.

Substantive Comment: My reconstruction puts this important point to good use.

> **262**: One might say: someone who has given himself a private explanation (German original, *Worterklärung*) of a word must inwardly *resolve* (German original, *vornehmen*) to use the word in such-and-such a way. And how does he resolve that? Should I assume that he invents the technique of applying the word; or that he found it ready-made

Translation Comment 1: Literally, 'Worterklärung' means 'word clarification.' It can also be translated as 'definition of a word.'

Translation Comment 2: A clearer translation of 'vornehmen' is 'to undertake.'

> **268**: And the same could be asked if a person had given himself a private explanation of a word; I mean, if he has said the word to himself and at the same time has directed his attention to a sensation.

Substantive Comment: Passage 268 also links up with Plato's Cave Allegory.

> **380**: I could not apply any rules to a *private* transition from what is seen to words. Here the rules really would hang in the air; for the institution of their application is lacking.
>
> **311**: "What greater difference can there be?" – In the case of pain, I believe that I can privately give myself an exhibition of the difference. – For the private exhibition, however, you don't have to give yourself actual pain; it is enough to *imagine* it. … This private exhibition is an illusion.

Substantive Comment: The type-token distinction is helpful in 311 as well. Imagining a sensation token without experiencing one is not like imaging having a sixth finger without actually having one, so in that sense the 'illusion' comment is correct.

> **580**: An 'inner process' stands in need of outward criteria (German original, *Ein 'innerer Vorgang' bedarf äußerer Kriterien*.)[16]

Grammatical Comment: 'Bedarf' is a (masculine) noun in German, so the word should have been written with a capital 'B' in the original because nouns are capitalized in German, e.g. 'Kriterion.' Perhaps it's a typo; perhaps the text was intended to read '*bedarft*' meaning 'requires.' Accordingly, I propose this:

> **580** (my translation): An 'inner process' requires outward criteria.

Substantive Comment 1: Passage 580 seems to imply that a person cannot decide whether he has followed a rule correctly using 'inner criteria,' which can at most result in a belief that he has followed a rule correctly. Whether the belief is true requires appeal to considerations that are 'external' to the belief. 580 seems to imply the correspondence theory of truth but I will not argue here that it does so.

Substantive Comment 2: Passage 580 should have been placed much earlier in the PI narrative, probably soon after 202. Stated as "A private process requires public criteria" 580 would have established a clear connection to other passages associated with a PLA.

[16] P.M.S. Hacker (1972, 277) agrees 580 is part of PLA, though he only provides a partial quote of the passage and does not identify it by number.

IX. Plato Anticipates

The Allegory of the Cave Plato described in Book VII of *Republic* comes remarkably close to anticipating the private language problem and the solution presented below. Here are the passages I have in mind in Tom Griffith's translation (Griffith 2000, 220-21.)

> ***Republic* 514a2-b8**: Picture human beings living in some sort of underground cave dwelling, with an entrance which is long, as wide as the cave, and open to the light. Here they live, *from earliest childhood*[17], with their legs and necks in chains, so that they have to stay where they are, looking only ahead of them, prevented by the chains from turning their heads. They have light from a distant fire, which is burning behind them and above them. Between the fire and the prisoners, at a higher level than them, is a path along which you must picture a low wall that has been built, like the screen which hides people when they are giving a puppet show, and above which they make the puppets appear.
>
> ***Republic* 515b6-8**: So if they were able to talk to one another, don't you think they'd believe that the things they were giving names to were the things they could see passing?[18]

[17] My italics. Griffith uses the locution "earliest childhood" in translating passages before and after 514a as well, e.g., 386a, 395c, 401d, 403d, 413c, 463d, 467d, 485d, 519a and 572c. Here are other translations:

- Shorey 1961, 747: "… legs and necks fettered from childhood."
- Grube 1974, 168: "The men have been there from childhood."
- Waterfield 1993, 240: "They've been there since childhood."
- Grube 1997 rev. Reeve, 1132: "They've been there since childhood."
- Reeve 2004, 208: "They have been there since childhood."
- Lee 2007, 241: "… since they were children ..."
- Rowe 2012, 239: "They have been there since childhood."
- Jowett 2016, 186: "Here they have been from their childhood."
- Bloom 2016, 193: "They are in it from childhood ..."

[18] Here are alternative translations of this passage, which do not form a consistent set:

- Shorey 1961, 747: "If then they were able to talk to one another, do you not think that they would suppose that in naming the things that they saw they were naming the passing objects?"
- Grube 1974, 168: "If they could converse with one another, do you not think that they would consider these shadows to be the real things?"
- Waterfield 1993, 241: "Now, suppose they were able to talk to one another: don't you think they'd assume that their words applied to what they saw passing by in front of them?"
- Grube 1997 rev. Reeve, 1133: "And if they could talk to one another, don't you think they'd suppose that the names they used applied to the things they see passing before them?"
- Reeve 2004, 208: "And if they could engage in discussion with one another, don't you think they would assume that the words they used applied to the things they see passing in front of them?"

What was the point of describing people in the Cave, now adults, as living in chains 'from earliest childhood,' or words to that effect. Why does it matter how long they've been living in such an appalling condition?

Assuming it does matter, there is an obvious explanation: To grab the reader's attention as per Drama 101. Every story has to have a hook, the more effective the hook, the better. Besides, Plato the dramatist was up against stiff competition. Sophocles, Aeschylus and Euripides were looking over his shoulder, not to mention Homer.

Because the author of *Republic* is a philosopher, however, the obvious won't do. I suggest we were told that people lived in chains 'from earliest childhood' so we would infer that language skills could not have been acquired through socialization, absent in the Cave. Having reached that conclusion, we were supposed to ask "How else could they have been acquired?" As it turns out, this is a trick question.

True enough, 515a tells us there were people in the Cave who milled about behind the prisoners, talking and carrying various objects whose shadows were projected on the wall. However, the narrative does not state or imply – which it could have – that they spoke to the prisoners. Considering that the prisoners' range of motion was restricted in a way that prevented them from turning their heads, it is clear they could not see or talk to the people carrying objects behind them. So that's that.

How, then, would prisoners have been able to acquire language skills sufficient for communicating with one another? Communication is ambiguous between act and content. The content part, according to Plato's doctrine of recollection,[19] was not acquired at all. It was present at birth. Like everyone else, the prisoners were born with a storehouse of shared meanings, which they could

- Lee 2007, 241: "Then if they were able to talk to each other, would they not assume that the shadows they saw were the real things?"
- Rowe 2012, 240: "So if the prisoners were able to have conversations with each other, don't you think they'd label whatever they were seeing in front of them as what those things actually are?"
- Jowett 2016, 186: "And if they were able to converse with one another, would they not suppose that they were naming what was actually before them?"
- Bloom 2016, 194: "If they were able to discuss things with one another, don't you believe they would hold that they are naming these things going by them that they see?"
- Zeyl 2022 (private communication): "So if they could converse with each other, don't you think that they would believe that the things they were seeing are the actual things?"

[19] I used to think this doctrine was outlandish until I struck stone with chisel for the first time in art class and had the feeling I had done it before, perhaps in a former life. I needed virtually no instruction in matters of technique. To this day I consider myself a kind of spectator when working on a sculpture. I try to stay out of the way and let my hands do as they wish. I'm at a loss how to explain this fact except anamnesis. This seems to me true in general of people considered 'naturals' at some activity.

draw upon to talk to one another once old enough to do so. At some point, how and when we do not know, prisoners would recollect the appropriate words and their meanings and use them to refer to the shadows on the Cave wall, including to identify and reidentify them as necessary.

So far, so good. Now, if we ask hypothetically what sort of meanings are possible in the Cave environment in the absence of socialization and a recollected storehouse of shared meanings, we arrive at the private language problem. Thus, the words prisoner X would use to refer to the images he sees on the wall can be understood only by X, likewise the words used by prisoners Y, Z and so on. Only private language games can be played in the Cave environment in such circumstances. But can they?

To connect with a PLA, with a bit of rewriting, passages 244, 257, 262 and 268 can be seen to apply to the Cave situation I've just described:

> **244C**: But how can the Cave prisoners set up the connection between the name and the thing named? This question is the same as: how did they learn the meaning of the names of sensations? – of the word 'pain' for example. Here is one possibility: The prisoners connected words with the primitive, the natural, expressions of the sensation and used it in their place.
>
> **257C**: So, do Cave prisoners understand a name, without being able to explain its meaning to anyone? When one says "He gave a name to his sensation", one forgets that much must be prepared in the language for mere naming to make sense [which are missing in the Cave.]
>
> **262C**: A Cave prisoner who has given himself a private explanation of a word must inwardly *resolve* to use the word in such-and-such a way. And how does he resolve that? Should we assume that he invents the technique of applying the word; or that he found it ready-made?
>
> **268C**: Suppose a Cave prisoner had given himself a private explanation of a word, e.g, 'blorse,' to be used to refer to a shadow on the wall resembling a horse; if he said the word 'blorse' to himself and at the same time directed his attention to the right shadow. [But, how does he decide which shadow is right and whether 'blorse,' if used again, was used correctly?]

Recall that, according to Plato, Cave prisoners can only hold beliefs, but can upgrade to knowledge once out of the Cave and in the sunlight of the Forms. As to how Cave prisoners can upgrade from belief to knowledge, Plato's answer is dialectic. What he meant by this concept is controversial (see Grube 1935, Friedländer 1958, White 1976, Gadamer 1980, Mueller 1992, Griffith 2000).

Here is Grube's view of the matter, which supports my second point that the Cave Allegory can be interpreted as anticipating my solution that there can be no such thing as a 'private dialectic,' which would never get the prisoners out of the Cave!

> **Grube 1935, 239**: As to the nature of this dialectic, it is clearly the power to think and express oneself logically. And as he has insisted that the objects of logical thought – the universals, the Forms – exist, he can speak of dialectic as the

> discovery of these Forms. For the present we may therefore consider dialectic mainly as the power of apprehending the Ideas, of thinking logically, so that the content of one's thought corresponds to Reality.

X. A Minimalist PLA Reconstruction

Here is the argument structure I propose, stated for sensation tokens. Steps 1-7 are premises; steps 8-13 follow by *modus tollens* as indicated. The next section fills in the details, including justification and links to PI.

1. If philosophical theory T is true, then, person P can refer to his own sensation tokens using words only he himself can understand.

2. Person P can refer to his own sensation tokens using words only he himself can understand only if person P can identify his own sensation tokens privately.

3. Person P can identify his own sensation tokens privately only if person P can reidentify his own sensation tokens privately.

4. Person P can reidentify his own sensation tokens privately only if person P can determine privately that sensation tokens he experienced at different times are tokens of the same type.

5. Person P can determine privately that sensation tokens he experienced at different times are tokens of the same type only if person P can follow a rule R privately to yield a valid argument for the proposition that sensation tokens he experienced at different times are tokens of the same type.

6. Person P can follow a rule R privately to yield a valid argument for the proposition that sensation tokens he experienced at different times are tokens of the same type only if person P following rule R privately to yield an argument for the proposition that sensation tokens he experienced at different times are tokens of the same type can satisfy condition C privately.

7. Person P following rule R privately to yield an argument for the proposition that sensation tokens he experienced at different times are tokens of the same type cannot satisfy condition C privately.

Therefore, by *modus tollens* from 7 and 6,

8. Person P cannot follow a rule R privately to yield a valid argument for the proposition that sensation tokens he experienced at different times are tokens of the same type.

Therefore, by modus *tollens* from 8 and 5,

9. Person P cannot determine privately that sensation tokens he experienced at different times are tokens of the same type.

Therefore, by *modus tollens* from 9 and 4,

10. Person P cannot reidentify his own sensation tokens privately.

Therefore, by *modus tollens* from 10 and 3,

11. Person P cannot identify his own sensation tokens privately.

Therefore, by *modus tollens* from 11 and 2,

12. Person P cannot refer to his own sensation tokens using words only he himself can understand.

Therefore, by *modus tollens* from 12 and 1,

13. Philosophical theory T is not true.

XI. Justifying the Premises

Premise 1: If philosophical theory T is true, person P can refer to his own sensation tokens using words only he himself can understand.

PI Sourcing for Premise 1:

> **243**: The words of this language are to refer to what only the speaker can know – to his immediate private sensations. So another person cannot understand the language.

PLA Targets: Traditional

> • The literature is in agreement that philosophical theory T can be attributed to Descartes and the British Empiricists. I will take it for granted this is true because it would take us too far afield to identify specific doctrines these philosophers held that assert or imply T; prove that T commits them to premise 1; and provide passages supporting these claims.[20] Nevertheless, it will be useful to cite two more recent potential targets of the PLA, Bertrand Russell and Rudolf Carnap.
>
> • Russell and Wittgenstein were friends and collaborators at one time (see Russell 1967-69; Clark 1975; and Monk 1996). Russell helped secure the publication of Wittgenstein's *Tractatus*, wrote a flattering review published in the book, and acted as examiner with G.E. Moore so Wittgenstein could get his PhD at Cambridge University. That all changed later. For example:

PLA Targets: Bertrand Russell

> **Russell 1959, 160-1**: Its positive doctrines seem to me trivial and its negative doctrines unfounded. I have not found in Wittgenstein's *Philosophical Investigations* anything that seemed to me interesting and I do not understand why a whole school finds important wisdom in its pages. … If it is true, philosophy is, at best, a slight help to lexicographers, and at worst, an idle tea-table amusement.

[20] In regard to Descartes, the reader may find it useful to have a look at Cusmariu 2021.

• Russell evidently failed to see that PLA took dead aim at views he had expressed decades earlier, for example in 1914. Here are relevant passages.

Russell 1914 [2019], 77: The first fact to notice is that, as far as can be discovered, no sensible[21] is ever a datum to two people at once. The things seen by two different people are often closely similar, so similar that the same *words* can be used to denote them, without which communication with others concerning sensible objects would be impossible. But, in spite of this similarity, it would seem that some difference always arises from difference in the point of view. Thus each person, so far as his sense-data are concerned, lives in a private world.

• Russell should have written 'no sensible token' rather than 'no sensible.' A sensible type can indeed be 'a datum to two people at once.'

• With a bit of work, which won't be attempted here, my PLA reconstruction can be extended to apply to Russell's views about sense-data. Passage 380 of PI hints as much:

380: I could not apply any rules to a *private* transition from what is seen to words. Here the rules really would hang in the air; for the institution of their application is lacking.

• For Russell, sense data are 'what is seen.'

• The PLA also applies to Russell's 1918 views:

Russell 1918 [1956], 198: A logically perfect language, if it could be constructed, would not only be intolerably prolix, but, as regards its vocabulary, would be very largely private to one speaker. That is to say, all the names that it would use would be private to that speaker and could not enter into the language of another speaker.[22]

PLA Targets: Rudolf Carnap

• Rudolf Carnap was also a one-time admirer of Wittgenstein's *Tractatus*, which became a sort of bible for 20th century logical positivists, for example:

• **Carnap 1934 [1995], 79-80**: In general, every statement in any person's protocol language would have sense for that person alone, would be fundamentally outside the understanding of other persons, without sense for them. Hence every person would have his own protocol language. Even when the same words and sentences occur in various protocol languages, their sense would be different, they could not even be compared.

Premise 2: Person P can refer to his own sensation tokens using words only he himself can understand only if person P can identify his own sensation tokens privately.

PI Sourcing for Premise 2:

256: Now, what about the language which describes my inner experiences and which only I myself can understand? How do I use words to signify (German original, *bezeichne*) my sensations? – As we ordinarily do? Then are my words for sensations tied up with my natural expressions of sensations? In that case

[21]What Russell calls 'sensibles,' other philosophers have called 'sensa,' 'sense data,' or 'qualia.'
[22] The discussion of Plato's Cave Allegory above bears on this issue.

> my language is not a 'private' one. Someone else might understand it as well as I. – But suppose I didn't have any natural expression of sensation, but only had sensations? And now I simply associate names with sensations, and use these names in descriptions.

- Premise 2 is true of non-privacy contexts as well. Its converse is false.
- Premise 2 seems true because its negation is unintuitive. That is, it would be odd in general for a person to be able to use the term 'toothache' to refer to his toothache but be unable to identify a sensation as a toothache rather than, for example, a backache. It seems that if he can do one, he can do the other.

Private identification of sensation tokens is a task that falls under schema PT.

> PT. Person P can perform task X privately =df Only P can understand how he himself is able to apply method M to perform task X.

Instantiating PT with respect to the identification of sensation tokens yields this:

> PTSTI. Person P can identify his own sensation token ST privately =df Only P can understand how he himself is able to apply method M to determine that ST exemplifies individuating property F.

- Informally, an individuating property is a property expressed by an open sentence of the form "*x* = the so-and-so" or "*x* = *a*," where '*a*' is a proper name or an individual constant.

Premise 3: Person P can identify his own sensations privately only if person P can reidentify his own sensations privately.

PI Sourcing for Premise 3

- The method M that is most consistent with the PI text is introspection, in this case a sort mental pointing described in PI variously. Recall that a PLA need not challenge the meaningfulness of introspection as a concept nor it application in the general case; only its application to carrying out certain tasks, in this case determining that a sensation token exemplifies an individuating property.
- This premise also true of non-privacy contexts. Its converse is true of privacy as well as non-privacy contexts.
- Assuming recurrence, identification implies reidentification. Passage 258 makes this point, where private methods of reidentification are mentioned and rejected.

> **258**: Let us imagine the following case. I want to keep a diary about the recurrence of a certain sensation. To this end I associate it with the sign 'S' and write this sign in a calendar for every day on which I have the sensation. -- I will remark first of all that a definition of the sign cannot be formulated. – But still I can give myself a kind of ostensive definition. – How? Can I point to the sensation? Not in the ordinary sense. But I speak, or write the sign down, and at the same time I concentrate my attention on the sensation – and so, as it were, point to it inwardly. – But what is this ceremony for? for that is all it seems to be! A

> definition surely serves to establish the meaning of a sign. – Well, that is done precisely by the concentrating of my attention; for in this way I impress on myself the connection between the sign and the sensation. – But "I impress it on myself" can only mean: this process brings it about that I remember the connection right in the future. But in the present case I have no criterion of correctness. One would like to say: whatever is going to seem right to me is right. And that only means that here we can't talk about 'right.'

Private reidentification of sensation tokens is also a task that falls under schema PT. Instantiating with respect to reidentification of sensation tokens yields this:

> PTSTRI. Person P can reidentify his own sensation token ST privately =df Only P can understand how he himself is able to apply method M to determine that ST re-exemplifies individuating property F.

Premise 4: Person P can reidentify his own sensations privately only if person P can determine privately that sensation tokens he experienced at different times are tokens of the same type.

- This premise is a near tautology. It is also true in non-privacy contexts.

Determining privately that sensation tokens experienced at different times are tokens of the same type also falls under schema PT. Instantiating yields this:

> PTSTD. Person P can determine privately that sensation tokens ST1 and ST2 experienced at different times are tokens of the same type T =df Only P can understand how he is able to apply method M to determine that sensation tokens ST1 and ST2 co-exemplify an individuating property of sensations of type T.

- It can be assumed without argument that sensation types can exemplify individuating properties as well as sensation tokens, though not necessarily the same individuating property.

Premise 5: Person P can determine privately that sensation tokens he experienced at different times are tokens of the same type only if person P can follow a rule R privately to yield a valid argument for the proposition that sensation tokens he experienced at different times are tokens of the same type.

- The doxastic burden imposed by Premise 5 isn't as heavy as it might appear. For example, the premise does not require person P to (a) determine anything privately; or (b) follow rule R privately; only that he can do so. Presumably, if he can do one, he can do the other.
- Following rule in this article means following a rule of inference, which in turn means applying the Substitution Theorem. Let us show that identification and reidentification are *modus ponens* arguments, so that rule R is *modus ponens*.

A Valid Token-type Identification Argument

> 1. If sensation token ST experienced at time t has property F, then sensation token ST is a token of sensation type S.

2. Sensation token ST experienced at time t has property F.

Therefore, by *modus ponens*

3. Sensation token ST is a token of sensation type S.

A Valid Reidentification Argument

1. If sensation token ST1 perceived at time t1 has property F and sensation token ST2 perceived at time t2 has the same property F, then sensation token ST1 and sensation token ST2 are tokens of the same type.

2. Sensation token ST1 perceived at time t1 has property F and sensation token ST2 is perceived at time t2 has the same property F.

Therefore, by *modus ponens*,

3. Sensation token ST1 perceived at time t1 and sensation token ST2 perceived at time t2 are tokens of the same type.

- Passage 244 suggests that property F is behavioral. (For my views on property identity, see Cusmariu 1978A.)

244: But how is the connection between the name and the thing named set up? This question is the same as: how does a human being learn the meaning of the names of sensations? – of the word 'pain' for example. Here is one possibility: words are connected with the primitive, the natural, expressions of the sensation and used in their place.

Following a rule privately is also a task that falls under PT. MP is the rule of interest here. Instantiating PT with respect to following MP privately yields this:

PTMP. Person P can follow MP privately to yield a valid argument for the proposition that sensation tokens he experienced at different times are tokens of the same type =df Only P can understand how he himself is able to apply MP to yield a valid argument for the proposition that sensation tokens he himself experienced at different times are tokens of the same type.

Premise 6: Person P can follow a rule R privately to yield a valid argument for the proposition that sensation tokens experienced at different times are tokens of the same type only if person P following rule R privately to yield an argument for the proposition that sensation tokens experienced at different times are tokens of the same type can satisfy condition C privately.

PI Sourcing for Premise 6:

- This premise applies Wittgenstein's dictum at 580 (my translation).

580: An 'inner process' requires outward criteria.

Condition C is the Substitution Theorem (ST). Satisfying ST privately is also a task that falls under PT. Instantiating and adding a suitable method yields this:

PTDT. Person P can satisfy ST privately = df Only P can understand P how he himself is able to satisfy the defining conditions of ST.

Premise 7: Person P following rule R privately to yield an argument for the proposition that sensation tokens experienced at different times are tokens of the same type cannot satisfy condition C privately.

PI Sourcing for Premise 7:

- Passage 259 of PI implies an affirmative answer to the question raised.
- **259**: Are the rules of the private language *impressions* of rules? – The balance on which impressions are weighed is not the *impression* of a balance.
- Premise 7 denies that only the person who is able to carry out uniform substitution in a valid argument form can understand how he himself is able to do so. Additional justification for this premise is provided in the next section.

XII. Questions and Answers

Question: Merely citing what comes across as an off-handed remark – PI is full of them – is not enough. Why shouldn't we conclude that your argument reached an impasse, unless further justification can be provided for Premise 7?

Answer: Premise 7 can be justified by means of a burden-of-proof argument (BPA).

- A BPA would challenge someone to explain how following rule R privately to yield an argument for the proposition that sensation tokens experienced at different times are tokens of the same type can satisfy condition C privately.

To get a sense of what is problematic about satisfying ST privately, here again is my informal explanation of what is involved in uniform substitution of propositional letters in MP, showing literally what is entailed.

(...) → (...)
(...)

(...)

How exactly would person P be able to follow the colors to fill in this structure such that that no one else could understand how P was able to do so?

- A BPA challenge would not stop at MP. Whoever thinks MP can be applied privately must hold that this can be done for any rule of inference in the quantificational as well as the propositional calculus. Good luck with that!
- But wait, there's worse! Whoever thinks there can be such thing as privately applied logic must explain how this can accomplish everything

standard logic can accomplish, including providing a foundation for mathematics and science.

- As co-author of *Principia Mathematica*, Russell would have vehemently denied that there can be such a thing as private applications of logic. It's a good question, then, why he didn't see that his views on privacy had this consequence if pushed far enough. This point also applies to Carnap.
- My PLA reconstruction shows we must choose between standard logic and being stuck inside the circle of our private ideas, never escaping Plato's Cave.

Question: Your argument proves a rather technical proposition, that it is not logically possible for person to make private use of language to refer to his own sensations, not the proposition that private languages as such are logically impossible. Isn't Wittgenstein after the more ambitious proposition?

Answer: The text favors a narrow interpretation of Wittgenstein's goals. A key passage, 243, seems to me to support the conclusion of my PLA reconstruction, namely, that privacy is to be applied to sensations and that 'the language' in this context is to be restricted to denoting phrases referring to sensations. Here is the passage once again:

> **243**: The words of this language are to refer to what only the speaker can know – to his immediate private sensations. So another person cannot understand the language.

Moreover, the term 'immediate' only makes sense here if it is applied to sensation tokens; that is, to sensations experienced by a specific person at a specific time and for a specific duration. The text is clear on restriction to denoting phrases referring to sensations.

Question: Passages 262 and 268 of PI describe private explanations of sensation words as involving a two part method: (a) uttering a sensation-type word to oneself and then (b) directing one's attention to a sensation token of that type. How would a PLA rule out private explanations of sensation words that followed this method?

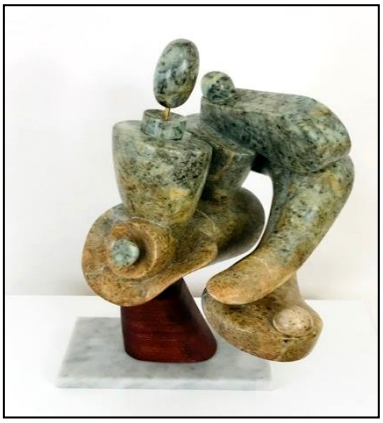
Answer: Step one would be to motivate the purpose of such a PLA by identifying an actual (rather than merely possible) philosophical theory and show that it was committed to the logical possibility of private explanations of sensation words that followed the method described in 262 and 268. Such motivation is not easy to find in the PI text, which often takes for granted that criticism of this or that method or solution is not aimed at a straw man, leaving it to others to fill in the blanks – a version of the Socratic Method, I suppose.

- Quite possibly Russell 1918 and Carnap 1934 accepted a theory that entailed the logical possibility of private explanations of sensation words described in 262 and 268. It also seems plausible to suppose that, absent socialization and a shared storehouse of remembered meanings, prisoners in Plato's Cave would be stuck having to learn the meaning of sensation words by following the method described in 262 and 268.

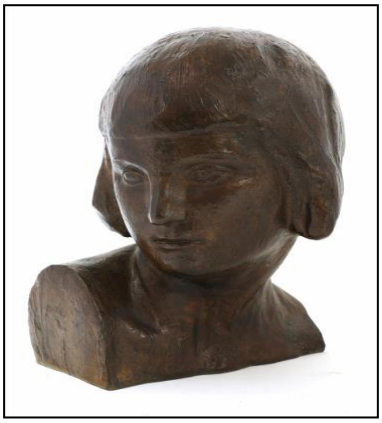
Step two would be to explain the sense in which explanations of the meaning of sensation words that followed the method described in 262 and 268 were private.

- To this point, privacy has been understood in epistemic terms. Thus, "person P applied out method M privately to perform task X" has been shorthand for "only P can understand his application of method M to his own performance of task X."
- What, then, would it mean for person P to apply conjunctive method M of uttering a sensation-type word to himself and then direct his attention to a sensation token of that type such that no one else could understand this application of method M?
- Perhaps a different concept of privacy would be appropriate. For example, "person P carried out method M privately" might be shorthand for "only person P can verify that he carried out method M." Using 'verify' would have a point, however, only if it meant "can verify that he had done it correctly." These are matters for another time.

XIII. Two Sculptures



Cusmariu: *Counterpoint A22*, 2019
Alabaster on mahogany and marble
16.5 inches high



Wittgenstein: *Head of a Girl*, 1925-28
Fired clay
15.5 inches high

- I noted earlier that my training in analytic philosophy and my views on ontology, which were and still are Platonist, significantly shaped my work as a sculptor. As I progressed in search of an individual voice in art – to 'make it new,' as Beethoven put it – I found it very helpful to give Plato's insightful metaphor of *interweaving forms* a musical interpretation, which led to the *Counterpoint* series.

- With that in mind, I thought readers might want to compare a sculpture by an analytic philosopher who is a Platonist with a sculpture by an analytic

philosopher whose sympathies were nominalist.[23] I will leave it to readers to study the two artworks and analyze differences. A few comments are in order.

- Wittgenstein's sculpture is a Renaissance knock-off. It doesn't rise to the level of creativity and imagination exemplified by the *Tractatus*. He doesn't seem to have realized that philosophical theories can have aesthetic impact at a practical level.
- Brâncuşi, Archipenko, Moore, Calder, Giacometti, and Lipchitz *inter alia* were doing something completely different at the time. I hope readers will see that my *Counterpoint A22* is completely different from what these major artists were doing a century ago. Cusmariu 2009, 2017A and 2017B explain my working aesthetic.
- Wittgenstein sculpture arguably illustrates the aesthetic poverty of nominalism, a philosophical theory opposed to abstract entities and by implication abstraction itself. While *Counterpoint A22* is a concrete object, as are all sculptures, its aesthetic attributes and mereology are abstract in a different sense of the term.
- Wittgenstein's sculpture is figurative – a copy of a copy – and as such runs afoul of Plato's objections to art in Book X of *Republic*. *Counterpoint A22* evades those objections.[24]

References

Ayer, A.J. 1966. "Can There Be a Private Language?" In *Wittgenstein: The Philosophical Investigations*, edited by George Pitcher, 251-266. New York: Anchor Books.

Baker, G.P. 1998. "The Private Language Argument." In *Language & Communication* 18: 325–56.

Candlish, Stewart and George Wrisley, "Private Language," In *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition), edited by Edward N. Zalta, URL = <https://plato.stanford.edu/archives/fall2019/entries/private-language/>.

Canfield, John V. 1975. "'I Know that I'm in Pain' Is Senseless." In *Analysis and Metaphysics: Essays in Honor of R.M. Chisholm*, edited by Keith Lehrer, 129-144. Dordercht: D. Reidel Publishing Company.

————. 2001. "Private Language: the Diary Case." *Australasian Journal of Philosophy* 79: 377–94.

Carnap, Rudolf. 1934 [1995]. *The Unity of Science*. Translated with an introduction by Max Black. Bristol: Thoemmes Press.

Castaneda, Hector-Neri. 1971. "The Private-Language Argument." In Jones: 214-239.

[23] Analytic philosophers Arthur Danto and Keith Lehrer also produced artworks – paintings, to be specific. I have not discerned any influence of analytic philosophy in their artworks, nor did they publish books or articles explaining their working aesthetic.

[24] My thanks Gary Rosenkrantz and Don Zeyl for helpful correspondence.

Chisholm, Roderick M. 1957. *Perceiving: A Philosophical Study*. Ithaca: Cornell University Press.

————.1966. *Theory of Knowledge*. Englewood Cliffs, NJ: Prentice Hall.

Clark, Ronald W. 1976. *The Life of Bertrand Russell*. New York: Knopf.

Cook, John W. "Wittgenstein on Privacy." In *The Private Language Argument*, edited by O.R. Jones, 240-272. London: Macmillan.

Crary, Alice and Rupert Read (eds.). 2000. *The New Wittgenstein*. London: Routledge.

Cusmariu, Arnold. Forthcoming 2023. *Logic for Kids: All Aboard the **Therefore** Train*. Singapore: Jenny Stanford Publishing.

————. 2021. "The *Cogito* Paradox." *Symposion*. *Theoretical and Applied Inquiries in Philosophy and Social Sciences* 8 (1): 7-43.

————. 2017A. "The Prometheus Challenge." *Symposion*. *Theoretical and Applied Inquiries in Philosophy and Social Sciences* 4 (1): 17-47.

————. 2017B. "The Prometheus Challenge Redux." *Symposion*. *Theoretical and Applied Inquiries in Philosophy and Social Sciences* 4 (2): 175-209.

————. 2016A. "Toward an Epistemology of Art." *Symposion. Theoretical and Applied Inquiries in Philosophy and Social Sciences* 3 (1): 37-64.

————. 2016B. "A Methodology for Teaching Logic-Based Skills to Mathematics Students." *Symposion*. *Theoretical and Applied Inquiries in Philosophy and Social Sciences* 3 (3): 259-292.

————. 2016C. "Semantic Epistemology Redux: Proof and Validity in Quantum Mechanics." *Logos & Episteme. An International Journal of Epistemology* VII (3): 287-304.

————. 2015A. "The Perils of Aphrodite: A New Take on Star Theory." *Film International* 13 (3): 97-116.

————. 2015B. "Baudelaire's Critique of Sculpture." *Journal of Aesthetic Education* 49 (3): 96-124.

————. 2012. "Toward a Semantic Approach in Epistemology." *Logos & Episteme. An International Journal of Epistemology* III.4: 531-543.

————. 2009. "The Structure of an Aesthetic Revolution." *Journal of Visual Arts Practice* 8 (3): 163-179.

————. 1985. "Self-Predication and the 'Third Man'." *Grazer Philosophische Studien* 23: 105-118.

————. 1980. "Ryle's Paradox and the Concept of Exemplification." *Grazer Philosophische Studien* 10: 65-71.

————. 1979A. "Russell's Paradox Re-Examined." *Erkenntnis* 14: 365-370.

————. 1979B. "On an Aristotelian Theory of Universals." *Australasian Journal of Philosophy* 57 (1): 51-58.

————. 1978A. "About Property Identity." *Auslegung* 5 (3): 139-146.

————. 1978B. "Self-Relations." *Southern Journal of Philosophy* 16 (4): 321-327.
Donagan, Alan. "Wittgenstein on Sensation." In *Wittgenstein: The Philosophical Investigations*, edited by George Pitcher, 324-351. New York: Anchor Books.
Dummett, Michael. "Wittgenstein's Philosophy of Mathematics." In *Wittgenstein: The Philosophical Investigations*, edited by George Pitcher, 420-447. New York: Anchor Books.
Fann, K.T. 1967. *Ludwig Wittgenstein: The Man and His Philosophy*. New York: Dell.
Friedländer, Paul. 1958. *Plato: An Introduction*. Translated by Hans Meyerhoff. Princeton: Princeton University Press.
Gadamer, H.-G. 1980. *Dialogue and Dialectic: Eight Hermeneutical Studies on Plato*. Translated by P. Christopher Smith. New Haven: Yale University Press.
Garver, Newton, 1971. "Wittgenstein on Private Languages." In *The Private Language Argument*, edited by O.R. Jones, 187-196. London: Macmillan.
Gellner, Ernest. 1959. *Words and Things*. London: Routledge.
Gram, Moltke. 1971. "Privacy and Language." In *The Private Language Argument*, edited by O.R. Jones, 298-327. London: Macmillan.
Grube, G.M.A. 1935. *Plato's Thought*. London: Methuen.
Hacker, P.M.S. 1972. *Insight and Illusion*. London: Oxford University Press.
Hahn, Lewis Edwin (ed.). 1977. *The Philosophy of Roderick Chisholm*. Chicago: Open Court.
Hardin, Clyde Laurence. 1971. "Wittgenstein on Private Languages." In *The Private Language Argument*, edited by O.R. Jones, 173-186. London: Macmillan..
Hymers, Michael. 2017. *Wittgenstein on Sensation and Perception*. London: Routledge.
Jones, O.R. (editor). 1971. *The Private Language Argument*. London: Macmillan.
Kenny, Anthony. 1966. "Cartesian Privacy." In *Wittgenstein: The Philosophical Investigations*, edited by George Pitcher, 352-370. New York: Anchor Books.
————. 1973. *Wittgenstein*. Cambridge: Harvard University Press.
Klemke, E.D. 1971. *Essays on Wittgenstein*. Chicago: University of Illinois Press.
Kripke, Saul A. 1982. *Wittgenstein on Rules and Private Language*, London: Blackwell.
Kraut, Richard (editor). 1992. *The Cambridge Companion to Plato*. Cambridge: Cambridge University Press.
LeBlanc, Hughes. 1966. *Techniques of Deductive Inference*. Englewood Cliffs NJ: Prentice Hall.
Malcolm, Norman. 1966. "Wittgenstein's *Philosophical Investigations*." In *Wittgenstein: The Philosophical Investigations*, edited by George Pitcher, 65-103. New York: Anchor Books.
Manser, Anthony. 1969. "Pain and Private Language." In *Studies in the Philosophy of Wittgenstein*, edited by Peter Winch, 166-183. London: Routledge.
McGinn, Colin. 1984. *Wittgenstein on Meaning*. Oxford: Blackwell.
McGinn, Marie. 1997. *Wittgenstein's Philosophical Investigations*. London: Routledge.

Monk, Ray. 1990. *Ludwig Wittgenstein: The Duty of Genius*. New York: The Free Press.

————. 1996. *Bertrand Russell: The Spirit of Solitude, 1872-1921*. New York: The Free Press.

Mueller, Ian. 1992. "Mathematical Method, Philosophical Truth". In *The Cambridge Companion to Plato*, edited by Richard Kraut, 170-199. Cambridge: Cambridge University Press.

Nielsen, Keld Stehr. 2008. *The Evolution of the Private Language Argument*. London: Routledge.

Peirce, Charles Sanders. 1906. "Prolegomena to an Apology for Pragmaticism." *Monist* 16: 492–546.

Pitcher, George. 1964. *The Philosophy of Wittgenstein*. Englewood Cliffs: Prentice Hall.

Pitcher, George (ed.). 1966. *Wittgenstein: The Philosophical Investigations*. New York: Anchor Books.

Plato. *Republic*. 1961. Translated by Paul Shorey. In *The Collected Dialogues of Plato*, edited by Edith Hamilton and Huntington Cairns, 575-844. Princeton: Princeton University Press.

————. *Republic*. 1974. Translated by G.M.A. Grube. Indianapolis: Hackett.

————. *Republic*. 1993. Translated by Robin Waterfield. Oxford: Oxford University Press.

————. *Republic*. 1997. Translated by G.M.A. Grube, rev. by C.D.C Reeve. In Plato, *Complete Works*, edited by John M. Cooper, 971-1223. Indianapolis: Hackett.

————. *Republic*. 2000. Translated by Tom Griffith. London: Cambridge.

————. *Republic*. 2004. Translated by C.D.C Reeve. Indianapolis: Hackett.

————. *Republic*. 2007. Translated by Desmond Lee. New York: Penguin Classics.

————. *Republic*. 2012. Translated by Christopher Rowe. New York: Penguin Classics.

————. *Republic*. 2016. Translated by Benjamin Jowett. Digireads.com Publishing.

————. *Republic*. 2016. Translated by Alan Bloom. New York: Basic Books.

Russell, Bertrand. 1914 [2019]. "The Relation of Sense Data to Physics." *Scientia* 16 (16): 1-27. Reprinted in Russell, Bertrand. *Mysticism and Logic and Other Essays*, 71-86. Dumfries & Galloway: Anodos Books. Page references are to this edition.

————. 1918 [1956]. "The Philosophy of Logical Atomism." Reprinted in *Logic and Knowledge*, edited by Robert C. Marsh, 178-281. London: Macmillan.

————. 1921. "Introduction" to Wittgenstein's *Tractatus-Logico Philosophicus*. London: Routledge.

————. 1959. *My Philosophical Development*. New York: Allen & Unwin.

————. 1967-69. *Autobiography*. Three volumes. London: Routledge.
Ryle, Gilbert. 1939. "Plato's 'Parmenides'." *Mind* 48: 129-151.
Schilpp, P.A. (ed.). 1959. *The Philosophy of C.D. Broad*. New York: Tudor.
Shwayder, D.S. 1969. "Wittgenstein on Mathematics." In *Studies in the Philosophy of Wittgenstein*, edited by Peter Winch, 66-116. London: Routledge.
Stern, David G. 2004. *Wittgenstein's Philosophical Investigations: An Introduction*. Cambridge: Cambridge University Press.
Strawson, P.F. 1966. "Review of Wittgenstein's *Philosophical Investigations*." In *Wittgenstein: The Philosophical Investigations*, edited by George Pitcher, 22-64. New York: Anchor Books.
Stroud, Barry. 1966. "Wittgenstein and Logical Necessity." In *Wittgenstein: The Philosophical Investigations*, edited by George Pitcher, 477-496. New York: Anchor Books.
————. 2000. *Meaning, Understanding, and Practice*. Oxford: Oxford University Press.
Todd, William. 1971. "Wittgenstein and Private Languages." In *The Private Language Argument*, edited by O.R. Jones, 197-213. London: Macmillan.
Tomas, Vincent. 1959. "Aesthetic Vision." *The Philosophical Review* 68 (1): 52-67.
White, Nicholas P. 1976. *Plato on Knowledge and Reality*. Indianapolis: Hackett.
Winch, Peter (ed.). 1969. *Studies in the Philosophy of Wittgenstein*. London: Routledge.
Wittgenstein, Ludwig. 1961 [1921]. *Tractatus-Logico Philosophicus*. Translated by D.F. Pears and B.F. McGuiness. London: Routledge.
————. 1958 [1953]. *Philosophical Investigations*, 3rd edition, translated by G.E.M. Anscombe. London: Blackwell.
————. 2009 [1953]. *Philosophical Investigations*, 4th edition, translated by G.E.M Anscombe, P.M.S. Hacker and Joachim Schulte. Revised by P.M.S. Hacker and Joachim Schulte. London: Blackwell. All translations are from the H&S edition unless otherwise indicated.
Wrisley, George. 2011. "Wittgenstein's Private Language Argument." In *Just the Arguments*, edited by Michael Bruce and Steven Barbone, 350-354. London: Blackwell.

Sen and Žižek on the One-dimensional View of Pathological Subjective Violence

Marlon Jesspher De Vera

Abstract: This paper presents an argument synthesized from the works of Sen and Žižek on how the one-dimensional view of pathological subjective violence is a mystification of the idea of violence. First, the paper provides an elaboration of the concept of objective violence as opposed to (but nonetheless still in relation to) subjective violence. Second, the paper follows with a discussion of the dialectics of the colonized mind as an example of how the objective violence of past colonialism is linked to the instigation of subjective violence even in recent times. Third, the paper provides a brief description of symbolic violence as another category of violence that is distinct from subjective violence. Lastly, the paper asserts its main argument on the mystification of subjective violence and proposes an alternative and more nuanced view of the mechanisms and causes of violence.

Keywords: Sen, Žižek, violence, subjective violence, objective violence, mystification.

The discussions in this paper draw from the works of Amartya Sen and Slavoj Žižek, and focus on the preoccupation of the conventional understanding of violence as pathological subjective violence, which is a mystification that pertains to the mechanisms and causes of violence. Žižek defines subjective violence as the type of violence enacted by a clearly identifiable social agent, whether this agent is an individual, a group, or an institution (2008b, 10). The term pathological here pertains to something that is caused by a malfunction or an unintended defect in an otherwise seamless mechanism. Thus, in the context of violence, pathological subjective violence is violence that is perpetuated by a clearly identifiable agent and which arose out of something that went wrong with the particular agent which resulted in the agent becoming violent. The discussions in this paper attempt to show how this one-dimensional view of violence is a mystification.

Objective Violence

The starting point of the discussions in this paper is Žižek's distinction between subjective violence and objective or systemic violence, as well as Žižek's discussion of how these two types or violence are inextricably linked. Subjective violence, on one hand, is the readily observable manifestation of violence through its disruptions of the normal, orderly, peaceful, and stable state of things. Objective violence, on the other hand, is not as readily perceived because it is precisely what makes the normal, orderly, peaceful, and stable state of things

possible. Of course, Žižek's Marxist analysis situates objective violence as a necessary element of the totality of the mechanism of global capitalism.

Carl Packman provides additional perspective on the character of objective violence in his essay *Towards a Violent Absolute: Some Reflections on Žižekian Theology and Violence* (2009). Packman posits that what makes the character of objective violence ambiguous is not that it is mostly hidden and absent, that it is seen only when probed and exposed. Rather, what makes it ambiguous is that it is too present and exposed in the normal state of affairs that the subject has gotten too desensitized to it and does not notice it unless there is a radical change to it. Packman makes a further connection between objective violence and how ideology functions in current times, and thus, in effect, explicates clearly the ideological dimension of objective violence. In the same way that the contemporary subject is offended only by explicit ideology and not by the implicit ideology embedded in the normal state of affairs that has become, so to speak, too obvious and too present to be noticed, the contemporary subject likewise focuses too narrowly on subjective violence and fails to see the objective violence that has also become too ingrained in the normal state of affairs to be noticed.

To further illustrate, Packman links the ideological dimension of the phenomenon of objective violence with two of Žižek's specific critiques in *Violence: Six Sideways Reflections* (2008b). First is Žižek's critique of what he calls 'liberal communists,' which are the extremely wealthy capitalists of the current time who present themselves as philanthropists first, businessmen second. They are able to do such ideological manipulation precisely because the contemporary subject perceives more strongly the explicitly presented ideology of benevolent charity and fails to readily appreciate the objectively violent underside of uncontrolled capitalism which has become, in effect, too obvious to be noticed. Second is Žižek's critique of Sam Harris' defense of torture through the thought experiment of the 'truth pill,' which is a hypothetical pill that, when taken by the subject to be tortured, would result in no external manifestations of suffering; instead, the extreme suffering of torture will only be manifested in the subject's internal experience. Žižek's critique of this thought experiment is that the idea of a 'truth pill' seems more tolerable than conventional torture only because it at least partially extinguishes the proximity of the tortured subject to the torturer. In such a way, the torturer would perceive an explicit ideology that is devoid of suffering, which would enable it to ignore the implicit but nonetheless all too obviously present suffering involved.

Furthermore, Packman situates Žižek's reflections on violence within what he calls the theological turn in Žižek's general philosophical project. This so-called theological turn is characterized by Žižek's explicit attempts to connect contemporary political analysis with their theological or religious philosophical roots, in many of his recent works since 1999. Packman argues that, although Žižek did not explicitly frame his analyses of violence in *Violence: Six Sideways Reflections* in theological or religious terms, he opened the door for further

explication through his numerous references to religious violence, divine violence, and the connections between violence and the notion of the neighbor or the other, to the point of asserting in one instance that ethics is the gap that separates Judaism and Christianity. My own reading is that theological and religious elements are definitely present in some of Žižek's reflections on violence, but that they only further enrich the comprehensiveness and complexity of the reflections by drawing attention to these important elements of the analysis.

But what is of primary interest to the discussions in this paper is Žižek's Hegelian assertion of how an excess of objective violence is necessarily accompanied by an excess of subjective violence. In other words, it is as if subjective violence is the inevitable consequence of objective violence (Žižek 2008b, 9-14). Immediately, it can be seen how this is a more expansive notion of the mechanisms and causes of violence that goes beyond the one-dimensional view of pathological subjective violence. Žižek's assertion can be linked to one possible caution against invoking the idea of objective violence, namely that the notion can make the conception of violence too broad to the dangerous extent that the idea of objective violence can already be used to justify the use of more subjective violence with the argument that it will be used to fight or undermine greater objective or systemic violence. In response to this caution, it can be argued that such invocations of the idea of objective violence to justify the use of more subjective violence is problematic when viewed through Žižek's account of objective violence. This is because, in Žižek's account, such use of more subjective violence is not a justified solution to greater objective or systemic violence. Rather, it is a constitutive symptom or indicator of the greater objective or systemic violence. Žižek's balanced response to such claims, to justify the use of more subjective violence through the idea of objective violence, is to object and fight against such claims, while not losing sight of the connection of such claims to the underlying objective or systemic violence that is present.

Sen does not make a direct similar discussion on the distinction between objective and subjective violence in his text but, nonetheless, he makes an analytical assessment of how certain features of the contemporary globalized market economy reinforce violence. The three examples that Sen cites, referring specifically to the persistent state of violence in Africa, are: 1) the continuous trade of arms from developed countries to African countries, which contributes to the perpetuation of totalitarianism and political militarism; 2) the structural injustices and inequitabilities of international policies on patents, which prevent inexpensive access to life-saving medicines, particularly for HIV; and 3) the highly restrictive export policies from developing countries to developed countries (Sen 2006, 95-96). Conceivably, Žižek would argue that these attributes of global capitalism in relation to Africa, which could be thought of as examples of covert underlying objective violence that ultimately gets manifested into subjective violence, are not contingent but rather necessary elements of the logic of capitalism and Africa's inclusion into its totality. Sen, on the other hand, argues

that structural changes can be made in the current globalized market economy, particularly in terms of revisions to relevant international policies, to radically change the current situation in Africa.

Sen makes a more elaborate assessment of the relationship between markets and freedom, particularly on how markets can both promote and undermine human freedoms, in his essay *Markets and Freedoms: Achievements and Limitations of the Market Mechanism in Promoting Individual Freedoms* (1993). Sen starts his assessment by reformulating the problem of evaluating the achievements and limitations of the competitive market mechanism beyond the conventional welfarist assessment, wherein the merits of the market economy are evaluated based on the achievement of individual and collective welfares rather than based on the promotion of human freedoms. Sen then invokes the distinction between the process aspect and the opportunity aspect of freedom and proceeds to raise the problem of evaluating the achievements and limitations of the competitive market mechanism in the context of each. In the realm of the process aspect of freedom, Sen acknowledges that the market mechanism indeed has merits in expanding the available processes of free choice for human beings and can thus be said to be enhancing process-freedoms. On the other hand, in the area of the opportunity aspect of freedom, the evaluation is more complex. While the competitive market economy can be instrumental towards the expansion of actual functioning opportunities that people can choose, value, and have reasons to value, and, consequently, can potentially expand human capabilities, the major problem that arises in the dynamics of opportunity-freedoms and the market economy is the problem of equitable distribution of substantive opportunity-freedoms among individuals. Thus, Sen identifies that the primary problematic limitation of the competitive market mechanism in promoting opportunity-freedoms has a lot to do with the problem of equity. Sen further asserts that while the problem of equity is already clearly pointed out in the conventional welfarist configuration of the problem of evaluating the merits of the competitive market mechanism, the problem of equity becomes more pronounced and compelling in the freedom-centric configuration. Sen's balanced assessment of the achievements and limitations of the competitive market mechanism in promoting human freedoms is consistent with his general position that recognizes how the market economy can potentially both undermine and perpetuate violence.

Evaluating the dynamics between the competitive market mechanism and human individual freedoms has been an ongoing long-term endeavor of inquiry for Sen. This is already apparent in Sen's earlier works, particularly in his 1970 landmark essay *The Impossibility of a Paretian Liberal* (1970). In this essay, Sen demonstrated that, at the basic level of logic, the ideals of social efficiency and individual freedoms are inherently in conflict with each other. Sen first defined the ideals of social efficiency and individual freedoms by formulating weak logical conditions of Pareto optimality and liberalism, respectively. He then formulated a logical theorem that, in a minimally sensible collective, no social decision function

can satisfy the minimal logical conditions of Pareto optimality and liberalism simultaneously, and subsequently provides a logical proof for the theorem. Although this logical conclusion understandably has limited immediate application, it has very important implications on the broader social and political problem of optimizing both social efficiency and individual liberty. It also greatly challenges the conventional assertion that the competitive market mechanism necessarily promotes individual freedoms as well. The basic position presented is that social decisions can be directed towards the promotion of either social efficiency or individual liberties, but never both, i.e., there will always be a trade-off between the two. The seminal ideas presented in this essay connect neatly with Sen's subsequent elaborative evaluations of the relationships between human individual freedoms and the competitive market mechanism.

The Dialectics of the Colonized Mind

Along a similar but distinct line of discussion, there is an important point of convergence between Sen and Žižek in their discussions on the central role of resentment and a sense of inferiority, particularly in the context of the colonial experience, as driving forces that instigate subjective violence. Based on my reading, both resentment and a sense of inferiority can be thought of as intermediate by-products of objective violence which catalyze the ultimate translation into subjective violence.

Žižek's analysis on resentment and a sense of inferiority focuses on so-called fundamentalists, and how it is compounded rather than undermined by the culturalist or politically correct approach of contemporary liberalism.

> The problem with fundamentalists is not that we consider them inferior to us, but rather that they themselves secretly consider themselves inferior. This is why our condescending, politically correct assurances that we feel no superiority towards them only make them more furious and feeds their resentment. The problem is not cultural difference (their effort to preserve their identity), but the opposite fact that the fundamentalists are already like us, that secretly they have already internalized our standards and measure themselves by them. (2008b, 86)

Sen follows the same line of analysis, but primarily focuses on how colonialism is linked with the mechanisms of violence, in an almost similarly Hegelian manner as Žižek, in characterizing and asserting that "the dialectics of the colonized mind includes both admiration and disaffection." (Sen 2006, 84) Sen then proceeds to develop this argument in a way that is rooted in his analysis of the mystifications of the idea of identity as manifested in identity-based thinking of colonized people. Sen refers to the sense of resentment and inferiority resulting from colonial humiliation (on top of the actual economic and political oppression imposed by the colonizer onto the colonized) as the 'reactive self-perception' of colonized people, which, when superimposed into identity-based thinking, results in identity-based alienation and its concrete negative repercussions.

> …the nature of this 'reactive self-perception' has had far-reaching effects on contemporary affairs. This includes (1) the encouragement it has given to needless hostility to many global ideas (such as democracy and personal liberty) under the mistaken impression that these are 'Western' ideas, (2) the contribution it has made to a distorted reading of the intellectual and scientific history of the world (including what is quintessentially 'Western' and what has mixed heritage), and (3) the support it has tended to give to the growth of religious fundamentalism and even to international terrorism. (2006, 89)

Thus, the dialectics of the colonized mind leads to violence when the sense of resentment and inferiority of the colonized is translated into a sort of identity-based thinking wherein the colonized defines its identity as essentially the other apart from the colonizer. Some examples that Sen cites are: 1) the Indian anticolonial nationalistic view, which purports a dualistic distinction between the West as material and India as spiritual; 2) the mistrust in Western science, which contributed to the spread of AIDS in Africa as posited by Mamphela Ramphele; and 3) the formulation of 'Asian values' and the 'Lee thesis,' to differentiate a more totalitarian and disciplined social and political system as distinctly Asian as compared to Western liberalism and democracy (2006, 90-93).

Sen further analyses how the 'reactive self-perception' of the colonized may not necessarily manifest immediately as violence. He makes an empirical observation on how times of extreme oppression, suffering, and poverty are usually accompanied by periods of peace and silence. Nonetheless, Sen asserts that the resentment and sense of inferiority will not be easily forgotten and instigators of violence can eventually instrumentalize the dialectics of the colonized mind for the purpose of violence (2006, 143).

Inferences from the analyses of both Sen and Žižek could likewise provide a potential characterization of the form of the violence that could arise either as the necessary compliment of objective violence or as the consequence of the dialectics of the colonized mind. Since both causes or mechanisms ultimately draw force from a depoliticized, particularist, and to some extent, irrational origin, it is likely that the resulting violence would also take a similar form. Žižek's analysis on what he calls violence as phatic communication shows this. The example that Žižek cites are the riots in suburban Paris in 2005, popularly presented as a form of racial and ethnic conflict (and thus, culturalized). However, upon closer reading, the riots could be seen as a mere explosion of irrational violence with no clear political agenda, clearly identified predicaments, or proposed solutions, but simply carried out to assert visibility, to demand to get noticed (Žižek 2008b, 76-77).

Symbolic Violence

Apart from subjective violence and objective or systemic violence, Žižek also identifies another category of violence which is symbolic violence, which is the violence in language itself, in its very symbolization of reality. Žižek makes a

Hegelian claim that the symbolization of reality itself is violent because the imposition of reality into the symbolic field of language is tantamount to reality's mortification (2008b, 61). Žižek, in his essay *Language, Violence and Non-Violence* (2008a), also discusses the paradoxical character of language as the locus of both violence and non-violence. While language is the medium of understanding, reconciliation, and convergence through discourse, it is also already inherently violent in its symbolization of reality. Thus, Žižek speculates that perhaps human beings have greater propensity towards violence precisely because of our capability to employ language (2008a). Despite its initial ambiguity, this notion could be an important starting point towards responding to the question Sen poses – why is the illusion of singular identity so effective in instigating violence (Sen 2006, 175)? Sen's motivation in posing this question is that, in his reckoning and most likely in common reckoning as well, the recognition that a person has multiple and diverse identities and that identity-based thinking must be accompanied by rationality and choice are both unremarkable recognitions. So how then do instigators of violence effectively make people ignore these unremarkable recognitions? Sen's response is that instigators of violence identify a singular identity affiliation that is part of the real identity of a person and then redefine the demands of this singular identity into a violent and belligerent form (Sen 2006, 176). Of course, this is a valid and sensible response, but I believe that this response itself could be explained further using Žižek's notion of symbolic violence. What the instigators of violence do could be seen as a violent symbolization of the reality of identity, imposing a narrow and extremely mortifying symbolic field on a person's identity-based thinking, such that the somewhat necessary consequence is the manifestation of this violently symbolized identity into its readily perceived form, which is subjective violence.

Conclusion

Given the lines of analysis and discussions presented in this paper on the mystification of the one-dimensional view of pathological subjective violence, some of the theses that can be arrived at are the following. First, the one-dimensional view of pathological subjective violence can be situated within the broader mystification of the culturalization and depoliticization of violence, as this view particularizes the pathology of the subject as the primary mechanism or cause of violence. Second, despite Sen's and Žižek's differences in terms of the philosophical traditions where they come from, it can be said that they both recognize that there are objective or systemic mechanisms or causes of violence beyond subjective pathology. Third, both Sen and Žižek recognize the significant force of the resentment and sense of inferiority of the oppressed, which can be thought of as an intermediate between objective violence and subjective violence. Lastly, the notion of symbolic violence could be employed towards a plausible explanation of the effectiveness of utilizing the illusion of singular identity to instigate subjective violence.

This paper elaborated on the complex and interlinked aspects of the mechanisms and causes of violence. The discussions in this paper challenged exclusively subject-centered conceptions of violence by introducing the notions of objective and symbolic violence. The particular mystifying notion that was critiqued in this paper is the idea that violence has to be recognized through the subject's perception and experience, and further that violence has to be some sort of intentional infliction by or undesirable intrusion to the subject. On the contrary, objective or systemic violence, in a seemingly paradoxical sense, is the kind of violence that perpetuates the normal and balanced status quo, which is already in itself violent, as it privileges a particular hegemonic political interest. Therefore, objective or systemic violence is often neglected not because it is mostly absent or hidden, but because it is too present and commonplace that the subjective view often fails to notice it.

The discussions on the dialectics of the colonized mind also showed how objective violence can potentially feed into a vicious cycle of resentment which, in a likely historically phased manner, could eventually lead to an explosion of subjective or other forms of violence. On the other hand, the discussions on symbolic violence attempted to more clearly characterize this seemingly ambiguous notion. Language, in its symbolization of the field of reality, is already inherently violent and it is this violence of symbolization that is often leveraged by instigators of violence to mystify notions of identity towards the instigation and perpetuation of violence. Nonetheless, language is also the most prominent locus of potential reconciliation and understanding through expression, compassion, and discourse. This paper also emphasized how the various potential forms, mechanisms, and causes of violence are inextricably linked in such a way that an excess in one form of violence would likely lead to an excess in the other forms of violence. This observation on the inextricable link among various potential forms, mechanisms, and causes of violence re-emphasizes the potent political dimension of the phenomenon of violence. Thus, these various forms, mechanisms, and causes must also be freely, rationally, and critically analyzed.

Lastly, this paper suggested that what can be considered as phenomena of violence can also be most likely considered as phenomena of unfreedom (considering a broad notion of freedom that takes both the process and opportunity aspects into account) or phenomena of irrationality (again, considering the broad sense of rationality or reasonability invoked in the discussions of this paper), and, vice versa, phenomena of unfreedom or rationality can also be most likely considered as phenomena of violence. Since freedom and rationality, broadly construed, both have objective/analytical and subjective/humanist components, such a conception of violence based on the promotion of freedom and rationality as the zero or base point can be analyzed in a more comprehensive manner across the various forms, mechanisms, and causes of violence – whether subjective (since, among other things, the subjective choices and actions of the agent, as well as the internal experiences and perceptions of the

subject, are both important considerations in the discourse of freedom and rationality), objective (since, among other things, the ample consideration of the objective and systemic social and political dimensions are constitutive to examinations of the process and opportunity aspects of freedom as well as of individual and collective rationality or reasonability), or symbolic (since, among other things, linguistic analyses are central to the analysis of freedom as capabilities and rationality through discourse).

References

Packman, Carl. 2009. "Towards a Violent Absolute: Some Reflections on Žižekian Theology and Violence." *International Journal of Žižek Studies* 3 (1).

Sen, Amartya. 1970. "The Impossibility of a Paretian Liberal." *The Journal of Political Economy* 78 (1): 152-157.

————. 1993. "Markets and Freedoms: Achievements and Limitations of the Market Mechanism in Promoting Individual Freedoms." *Oxford Economic Papers* 45 (4): 519-541.

————. 2006. *Identity and Violence: The Illusion of Destiny*. New York: W.W. Norton & Company.

Žižek, Slavoj. 2008a. "Language, Violence and Non-Violence." *International Journal of Žižek Studies* 2 (3): 1-12.

————. 2008b. *Violence: Six Sideways Reflections*. New York: Picador.

Should the State Teach Ethics? A Schematism

Landon Frim

Abstract: Should the state teach ethics? There is widespread disagreement on whether (and how) secular states should be in the business of promoting a particular moral viewpoint. This article attempts to schematize, and evaluate, these stances. It does so by posing three, simple questions:
(1) Should the state explicitly promote certain ethical values over others?
(2) Should the state have ultimate justifications for the values it promotes?
(3) Should the state compel its citizens to accept these ultimate justifications?
Logically, each question in this series is a prerequisite for considering those questions further down the list. The result is that responses can be categorized into one of four possible permutations or 'camps.' These are:
(1) The Libertarian ("No" to all three questions)
(2) The Pluralist ("Yes" to question 1; "No" to questions 2 and 3)
(3) The Rationalist Republican ("Yes" to questions 1 and 2; "No" to question 3)
(4) The Rigorous Republican ("Yes" to all three questions)
It will be shown that just one of these positions, the 'rationalist republican,' stands out from all the rest. For only the rationalist republican can account for a normative politics while also safeguarding the individual's freedom of conscience.

Keywords: Secularism, Republicanism, Toleration, Political Philosophy, Ethics, Baruch Spinoza.

Confessional states – those with an explicit, religious identity – teach ethics as a matter of course. Moral norms may even be enforced through civil and criminal penalties. A clear example of this is Afghanistan's "Ministry for the Propagation of Virtue and the Prevention of Vice," re-established in 2021 after the fall of Kabul to Taliban forces. (Pal 2021) By contrast, secular states are often loath to admit that their laws emanate from some coherent philosophy or ethical theory. It's often thought that the mark of a modern, secular state is an official 'neutrality' toward such speculative questions; these are instead the proper domain of private conscience. Nonetheless, it is easy to point out the innumerable ways in which secular states do, in fact, engage in systematic ethical training.

This is particularly evident in publicly funded schools. Students in the United States are often exposed to 'anti-bullying' campaigns endorsed by their respective districts, as well as programs to warn against premarital sex, recreational drug use, and to promote the qualities of 'grit,' environmentalism, and positive body image, to name just a few examples. (See, for example, the U.S. Department of Health and Human Services' "Bullying Prevention Training Center" n.d.) In Palm Beach County, Florida, an entire section of the curriculum has long

been dedicated to a study of the Holocaust and genocide – not merely as a dry recounting of historical facts, but rather as a normative attempt to warn against the dangers of racial hatred, xenophobia, discrimination, and ultra-nationalism. Such pedagogical approaches are mirrored in other school districts around the country.[1] (Staff Report 2021)

Adults are not exempt. Public sector employees, at the federal, state, and municipal levels, are routinely mandated to attend training sessions on the topics of workplace sexual harassment, the reasonable accommodation of people with disabilities, and workplace discrimination. ("U.S. Equal Employment Opportunity Commission" n.d.) In each of these cases, what at first appears to be the merely neutral instruction of positive law, good manners, hygiene, health, or decency, turns out (upon the most cursory inspection) to be grounded on a rather specific set of moral values and norms.

That all people are deserving of equitable treatment is a moral claim which may well be defended in any number of ways; It is not an uncontroversial, empirically obvious fact. That 'the good life' is marked by the values of toleration, egalitarianism, non-discrimination, grit, adolescent chastity, or sobriety are equally not self-evident propositions. We may wish to affirm some or all of these virtues as legitimate – even crucial – but to do so honestly means confronting what they in fact are: ethical norms. Likewise, government-endorsed programs which enshrine and promote such norms should be understood as instances of state-sponsored ethical instruction. And so, the question is not *whether* secular states engage in ethical indoctrination (they do), but rather *ought they*, and if so, then in *what manner*?

Cast in these stark terms, such a question will provoke innumerable, nuanced responses. At the risk of oversimplification, a schematism may be useful for sorting the majority of these, and quite possibly, clarifying the strengths and weaknesses of each. There are three essential questions which a political philosophy must answer regarding the public instruction of morals. These are:

(1) Should the state explicitly promote certain ethical values over others?

(2) Should the state have *ultimate justifications* for the values it promotes?

(3) Should the state compel its citizens to accept these ultimate justifications?

Each question in this schematism can be seen as a 'gate' or conceptual hurdle, such that answering no to an earlier question (e.g., "Should the state explicitly promote certain ethical values over others?"), logically precludes considering the latter questions (e.g., "Should the state have *ultimate justifications* for the values it promotes?"). The result of this schematism is that it sorts political

[1] Palm Beach County's "Holocaust Education Proclamation" explicitly uses the normative language of "the moral responsibilities of individuals, societies, and governments," remaining "vigilant against hatred, persecution, and tyranny," and that the district should "actively rededicate itself to the principles of individual freedom in a just society."

philosophies into one of four 'camps.' For the sake of clarity, we can tag these camps with the following designations:

The Libertarian ("No" to all three questions)

The Pluralist ("Yes" to question 1; "No" to questions 2 and 3)

The Rationalist Republican ("Yes" to questions 1 and 2; "No" to question 3)

The Rigorous Republican ("Yes" to all three questions)

The remainder of this paper will consist of delineating these four permutations and evaluating their relative strengths and weaknesses. It will be shown that just one of these positions, the 'rationalist republican,' stands out from all the rest. For only the rationalist republican can account for a normative politics while also safeguarding the individual's freedom of conscience.

The Libertarian ("No" to question 1)

Like the other designations used in this paper, 'libertarian' is here employed as a convenient label for those answering our three questions in a given way (in this case, replying "no" to each). That is, the state should not be in the business of promoting, justifying, or enforcing, an ethical doctrine. This is *not* meant to be a categorical statement about every libertarian thinker, and their total rejection of public moralizing as such.

To be sure, there are some self-identified libertarians, in the model of Andy Barnett, who explicitly seek to ground political libertarianism on an ethical basis. (Barnett 2004) Ayn Rand, still perhaps the most recognizable libertarian thinker, wrote that the fundamental achievement of American democracy was, "the subordination of society to moral law." Here, Rand had in mind the principle of individual rights, and specifically, the moral notion (common to Immanuel Kant) of "man as an end in himself."[2] (Rand 1964, 109)

Nevertheless, the tendency of libertarians to balk at the official promotion of morals is well-documented. Again, we can cite Ayn Rand in advocating for the "virtue of selfishness" as against the moralizing political agendas of both the Right and Left. Her individualist egoism (including a hedonistic view of sex and marriage and the right to contraception and abortion) flies in the face of the anti-choice Religious Right. (Rand 1990, 54–55) Similarly, her brand of individual self-reliance rejects what libertarians characterize as liberal social engineering, 'enforced mediocrity,' and economic levelling.

For Rand, traditional religious morality and modern liberalism are something of a piece – each promoting a form of slavish self-sacrifice of the individual before a faceless other (whether this be the State or God). Governments

[2] It is ironic that Rand would adopt such a Kantian moral formula given that she appears to have despised his philosophy, even referring to Kant as "the most evil man in history." (Kelley 2001, 57)

should not be in the business of promoting such a false virtue, either indirectly through social programs, or through explicit pedagogy in the school or workplace.

Retired congressman Ron Paul, famous for his libertarian politics, strikes a similar chord. In speaking about the ideal school curriculum, he insists that it "…never puts ideological indoctrination ahead of education." (Paul 2018) Though, coming from a personally religious background – including opposition to abortion and same-sex marriage – Paul is noticeably more accepting of the exposition of Judeo-Christian values than was Ayn Rand.

> Unlike government schools, and even many private schools, my curriculum addresses the crucial role religion played in the development of Western civilization. However, the materials are drafted in such a way that parents of any or no religious belief can feel comfortable using the curriculum. (Paul 2018)

Such thinking is not an aberration in contemporary libertarian thought, but is rather one of its common features. Socially-conservative libertarians, like Paul, tend to see public institutions as hostile to their sincerely held religious and moral beliefs. Their remedy is to remove all hints of liberal ideological bias within the curriculum, while 'not censoring' the important role that religion has played in American history and that of Western civilization at large. In other words, the state must be morally neutral so as not to conflict with the private, often religious moralities of its citizens.

Of course, a more direct strategy (also employed by Paul) is to devolve the role of education to parents, churches, and other private organizations entirely so that individuals have a choice as to where their children will learn, and what sort of values they will be imparted. Similarly, Right-libertarians may oppose federal bans on abortion and same-sex marriage; But this is only so that the authority to enact such restrictions can "devolve" to the various states or religious institutions. (Weiner 2011)

Thus, Paul promotes what he calls 'Ed-Exit,' (ostensibly a play on the portmanteau 'Brexit'). Federally-governed public education should be dismantled in the face of intractable cultural, religious, and ethical divides endemic to American society.

> Centralizing education at the state or, worse, federal level inevitably leads to political conflicts over issues ranging from whether students should be allowed to pray on school grounds, to what should be the curriculum, to what food should be served in the cafeteria, to who should be allowed to use which bathroom. (Paul 2016)

Note the striking formalism, and apparent ethical skepticism, at work here: Cafeteria food selection is put on moral par with church/state separation and discrimination against transgender students. Each is cast as just another irresolvable debate amongst a culturally-divided American populace.[3] Still, Paul is

[3] Dietary concerns can, of course, intersect with moral or spiritual beliefs. However, it appears unreasonable to equate offering a vegan, kosher, or halal cafeteria option to omitting slavery

no moral skeptic. His merely rhetorical formalism does the work of pushing all ethical questions out of the political realm and into the (typically more traditionalist) hands of private institutions and families.

This 'devolution' strategy is shared by certain neo-confederate strains of American libertarianism. This includes some who have worked directly under Ron Paul, notably Murray Rothbard and Llewellyn Rockwell Jr., founder of the Ludwig von Mises Institute. The latter has been widely reported as the ghostwriter for numerous homophobic and racist pieces in Paul's political newsletter. (Sanchez and Weigel 2008)

The push for devolution is also evident in the Right-libertarian works of Nick Land and Curtis Yarvin (a.k.a., Mencius Moldbug). Land's essay *The Dark Enlightenment* (which draws heavily on Yarvin's own thought), advances a scathing criticism of "PC culture," the welfare state, and liberal leveling. (Land 2012) This is libertarianism taken to its (neo)reactionary limits – where the ideal government is reconceived as a private corporation, and citizens are given the right of "free exit," but "no voice." (Mass democracy being hazardous to individual freedom.)

The liberal welfare state – what they derisively term "The Cathedral" – is a drag on progress. Its incessant need to regulate business, and to promote equity and human wellbeing, results in the destruction of vital "feedback loops." (Land 2012) Specifically, the nanny state intervenes to protect the inept, the incompetent, and the mediocre from suffering the market consequences of their own failures (including unemployment and poverty). Values-education in schools, and workplace seminars countering harassment and discrimination, are just so many examples of this counterproductive – and moralistic – meddling.

It is not clear to what degree libertarian thought has been infected by the authoritarian and chauvinist leanings of Rockwell, Rothbard, Land, or Yarvin. What does seem to be a libertarian consensus, however, is a general despair that the state can effectively promote unifying social values. Such attempts are seen as not only futile, but often as leading to social dissension and outright conflict.

The Cato Institute website, as if to drive the point home, hosts a "Public Schooling Battle Map" to showcase the innumerable skirmishes which arise when public institutions find themselves in the messy business of values education. Cato's take-away seems to be that values are nothing more than a zero-sum contest with no resolution in sight. As their website reads:

> Americans are diverse – ethnically, religiously, ideologically – but all must pay for public schools. The intention is good: to bring people together and foster social harmony. But rather than build bridges, public schooling often forces people into wrenching conflict…Think creationism versus evolution, or assigned readings containing racial slurs. The conflicts are often intensely personal, and

from the American history curriculum. The former merely accommodates individual student preferences while the latter involves a normative stance taken by the institution itself.

> guarantee if one fundamental value wins, another loses. ("Public Schooling Battle Map" n.d.)

Again, the entire problem is cast in highly formalistic and skeptical terms, as though moral divisions are simply irresolvable, brute facts of life. On this view, moral contests are not about the truth (where one side may be rationally convinced to change their position), but only about power – i.e., which political faction or identity group has captured the public institution, and so can foist their views on all others. This pessimism about reason and democracy is encapsulated by one of Søren Kierkegaard's favored maxims: "As soon as it has come to the point that the crowd is to judge what is truth, it will not be long before decisions are made with fists." (Kierkegaard 1997, 90 citing Friedrich Schelling)

Conceived in this way, libertarians are apt to vote for 'exit,' that is, the devolution of federal institutions to the states and private hands. For there is no telling when those hostile to your basic beliefs will come to power. As Jesse Walker of *Reason Magazine* sardonically put it, "Apparently, public institutions created with one set of values in mind can be captured by people with a different set of values. You may want to bear this in mind the next time you find yourself creating a public institution." (Walker 2012) But if we must live with some public institutions, reasons the libertarian, better that they not be empowered to promote certain values over others; Better, in other words, to answer "no" to our first question.

Nonetheless, we may ask whether such libertarian pessimism is warranted? Are divisions over fundamental values truly irresolvable? It appears that moral progress is indeed achievable, just so long as we consider massively changing attitudes about slavery, miscegenation, civil rights, and gender equality to count as progress. (Newport 2013)

Furthermore, it appears that a major driver of these changing cultural attitudes is their inclusion in the political realm of democratic deliberation. In other words, the "Breitbart Doctrine" that "politics is downstream from culture" is very often wrong. (Meyers 2011) Legislation, elections, and even ballot initiatives can act as focal points which galvanize communities to action, engender social consensus, and even generate new cultural identities. At the very least, there appears to be a dialectical relationship between cultural norms and political movements, rather than this being a one-way street (or 'stream').

This fact illuminates a central tension of the libertarian position: Government is simultaneously considered to be inept at promoting unifying social values, and at the same time, supremely threatening. A rhetorical skepticism about morals suggests that ethical divisions are irresolvable and undecidable in the public sphere. Common reason and deliberation can't select between contending values and so only brute power can impose these through coercion.

But as we have seen, such skepticism about moral truths is most often performative rather than a sincerely held belief. From Rand's "virtue of selfishness" to Paul's social conservatism, to Land and Yarvin's neoreactionary ethos,

libertarian politics is shot through with value statements. Only these values are never held up to democratic scrutiny. Instead, politics is neutralized so that private institutions, especially corporations, the family, and organized religion can impose their norms on their respective members.

This, of course, in no way makes values-education any less coercive or punitive. Children who are raised in a particular family or attend a parochial school will be subject to the ethos of those private and hierarchical institutions. So too will adults who are financially dependent on private-sector jobs or religious charities. Nonconformity will be met with social censure and real economic consequences. Thus, what is truly missing in the libertarian scheme is neither a conception of morality, nor mechanisms of coercion, but only democratic responsibility. For the selection and promotion of norms is placed, not in the hands of publicly elected bodies or a voting citizenry, but rather in the unaccountable hands of the rich, influential, and powerful.

Stepping back, we can take another lesson from the libertarian's *merely performative* moral skepticism: One cannot do without some conception of the good. Without this, political activity seems to lack sufficient motivation. In the absence of guiding values, why bother to vote or legislate at all? But if this is true, then how can we meaningfully divorce civic education from moral education? Our conception of 'the good' will always color our notion of what makes for a good law, a good republic, and a good citizen.

To deny that the state should promote certain values over others is thus problematic; For a state's values will regularly manifest themselves through common ordinances, budget priorities, and foreign policy agendas. Values-education only makes explicit (and thus open to public scrutiny, critique, and amendment) a state's guiding norms.

The Pluralist ("Yes" to question 1; "No" to questions 2 and 3)

Suppose, then, that we answer "yes" to question 1; The state *should* promote certain ethical values over others. This opens the door for considering a second query: "Should the state have *ultimate justifications* for the values it promotes?"

For our purposes, 'the pluralist' is that person who answers yes to question 1, but no to question 2. This is to say, the pluralist admits that the state inevitably affirms certain norms, and yet, should not try to ground these ideals in one, all-encompassing view of reality. There are no 'ultimate justifications,' or in other words, there is no 'state philosophy.'

If we go down the road of ultimate justifications, the pluralist surmises, we risk turning the state into a metaphysical schoolmaster – not only protecting citizens' rights and welfare, but also insisting on a particular worldview from which these goods are derived. This is simply not the role of government. Worse still, professing a very specific view of existence is bound to exclude (and perhaps even provoke) wide swaths of the populace who otherwise might happily sign on

to a more pragmatic moral consensus. We should respect that diverse citizens will affirm common laws and norms for equally diverse reasons.

The 'pluralist' category is as varied as its name suggests. It is comprised of modern liberals (including John Rawls, Isaiah Berlin, and Richard Rorty), as well as communitarian and post-secular critics of liberalism (especially Charles Taylor, Jürgen Habermas, and Étienne Balibar). But despite their substantive differences, what defines this class is a claim that sweeping, grand narratives are bad for politics.

In Rawls' work, such well-defined, grand narratives are given the name 'comprehensive doctrines.' In his view, it's not the role of the state to uphold up any one comprehensive doctrine as the ultimate truth. Instead, each citizen or community should be free to develop their own, diverse belief systems. Sincerely held beliefs about God, salvation, morality, and other ultimate questions, are a matter of personal conviction and not public policy. At the same time, when it comes to *public* decision making, the diversity of private opinions will be subordinated to some 'reasonable' liberal consensus. Public law and private conviction need not infringe upon one another's proper domains. These constitute, to borrow a phrase, "non-overlapping magisteria." (Gould 1999, 5)

This word, 'reasonable,' is of course very important for the Rawlsian pluralist. It denotes that sort of person who is willing to prioritize a tolerant and rights-based politics above their own particular belief system (especially when these conflict). The political freedoms of speech and association will supersede, for example, religious proscriptions against blasphemy and false worship.

The 'reasonable' Jew, Christian, Muslim, Hindu, or Atheist will, moreover, not try to base public law upon their privately-held metaphysics. Final views on truth, salvation, existence, and the good life are notoriously difficult to determine for oneself, let alone to convince a whole body politic. The continued debates amongst philosophers, and the whole bloody history of religious wars, seem to be *prima facie* evidence of this difficulty, what Rawls terms the "burdens of judgment." (Rawls 2005, 54–58)

But if one's comprehensive doctrines are subordinated to public law, they are not thereby expelled from the public domain altogether. This is what distinguishes the liberal pluralist from the libertarian. Instead, reasonable people of all stripes will find within their respective belief systems the motives for affirming core liberal values. That is, the reasonable Catholic will find Catholic reasons for defending the freedom of conscience and expression (perhaps a religious doctrine of free will); so too will the reasonable atheist (though on a totally different, secular basis). Social stability depends upon such an "overlapping consensus" of diverse individuals professing identical norms, but for non-identical reasons. (Rawls 2005, 150–209)

When it comes to the state teaching of ethics, then, the Rawlsian pluralist will offer a different sort of response as compared to the libertarian. For the libertarian, the state should not be in the business of professing moral norms at

all – that's what families, churches, charities, and corporations are for. For the liberal pluralist, however, the state should profess and promote the fundamental liberal norms (freedom of thought and expression first among them). Only, it should not profess the *basis* for these values because (1) such metaphysical questions are hard to resolve, and relatedly, (2) each reasonable citizen will defend these norms according to their own, diverse doctrines.

Churches, temples, and mosques are therefore not the places to challenge or displace public morality, but instead, to support it using their respective, sacred vocabularies. School curricula and workplace regulations, likewise, should promote the reasonable norms of anti-racism, respect for people with disabilities, and gender equality; Only, they must not uphold some *ultimate rationale* for such values (secular or otherwise).

Plausible as all this may sound, liberal pluralism quickly deconstructs itself. Rawls' notion of the "burdens of judgment" is not merely an empirical claim about how difficult or acrimonious philosophy can be. Rather, such pluralism relies upon an extreme anti-foundationalism. In Rawls' words, "…the correct regulative principle for anything depends on the nature of that thing..." (Rawls 2003, 25) There can be no grand, unified theory of everything. Instead, each human endeavor is subject to its own regulating principles. Politics is no exception.

The best we can do is to aim for a "reflective equilibrium" wherein general principles do not govern, but rather are reciprocal to, specific political judgments. (Rawls 2005, 8, 28, 45, 72) That women should have the right to vote, for example, is not derived from the prior, speculative belief that all people have equal inherent worth. Instead, this particular judgment about voting supports the general principle, *and vice versa*.

All specific judgments are fallible, given new evidence and circumstances. So too, therefore, are all the general principles which mutually depend on them. In this way, politics is functionally prior to political theory. Speculation cannot conceive of the good state on its own; Instead, the continued dialogue, debate, and justificatory processes of a democracy provide the real content for our political and moral ideals. (Rawls 1999, 286–302) An ethos of epistemic humility suffuses such liberalism, hesitating to submit unique, lived experiences to the Procrustean bed of theory.

Other liberal pluralists, notably Isaiah Berlin, pushed the envelope further: If foundationalism is false, then there can be no singular view of reality which grounds all liberal norms. And if this be the case, then individual liberal values may sometimes conflict, or even stand as incommensurable to one another. "…Great Goods can collide…some of them cannot live together, even though others can…one cannot have everything, in principle as well as in practice…" (Berlin 2013, 17)

Freedom and equality, equal opportunity, meritocracy, and basic welfare are distinct sorts of goods which no singular algorithm can adjudicate; "realising ultimate harmony is a fallacy..." (Berlin 2013, 17) For in this world, "we are faced

with choices between ends equally ultimate, and claims equally absolute, the realisation of some of which must inevitably involve the sacrifice of others." (Cherniss and Hardy 2022) Philosophy cannot decide between competing goods and rights. For this we must rely on our particular, lived experiences.

To the theoretically-inclined, all of this may sound rather threatening, as though the foundation for politics has turned to sand. Yet Berlin saw this plurality of competing values as a feature, not a bug, of modern politics. It mandates a process of open debate and discussion, rather than the high-handed rule of experts, theologians, and philosopher-kings.

Still, when the only arbiter between conflicting values is the concrete situation, then Rawls' "reflective equilibrium" rapidly morphs into Berlin's "precarious equilibrium." (Berlin 2013, 18–19) The internal stability of liberalism gets called into question, and with this, its supposed reasonableness. If liberal norms of freedom, equality, and toleration do not have a common grounding – if speculative 'grounding' is an anathema to begin with – then why should *any* of these norms supersede a citizen's private, comprehensive doctrines?[4]

Why should we presume that an ancient Muslim, Hindu, or Jewish doctrine is any less obvious, less fundamental, or more subject to the 'burdens of judgment,' than the relatively young, mutually-conflicting (and presently contested) norms of liberal pluralism? At a time when political liberalism is under sustained attack from religious extremists, populist demagogues, and ultra-nationalists, why should we imagine liberalism to be a point of 'reasonable' consensus for all people?

At its most consistent, liberal pluralism declines to justify itself at all. It accepts the sheer plurality, contingency, and historicity of all belief systems whatsoever (including its own). There is no such thing as a justified, true belief – at least in any non-circular sense. In pursuing this line of reasoning, Richard Rorty sketches a figure he calls the "liberal ironist" – liberal, because they believe that "cruelty is the worst thing we do," but at the same time, ironic, in facing up to "the contingency of his or her most central beliefs and desires…" (Rorty 1989, xv, 91)

This heroic figure sanguinely denies that there can be any privileged, "single vision" which stands "outside the various vocabularies we have used…" to do politics. But given this, they declare for an inclusive liberalism anyway. Human solidarity is not something we discover in Nature; It is not something "true" to be deduced by philosophers. Instead, it's an "achievement" we choose to pursue. (Rorty 1989, xvi, 77)

And why should we work for such an utterly contingent achievement? For consistent pluralists like Rorty, this is simply an unforced decision on our part. The era of philosophical justification has ended, much like the era of theological justification. (Rorty never tires of equating the two.) In his words, "A

[4] Berlin, like Rawls, attempts to answer this question through appeals to "the political culture." Yet this reasoning is circular seeing as the political culture may, itself, be steeped in faulty principles. (Berlin 2013, 18–19)

postmetaphysical culture seems to me no more impossible than a postreligious one, and equally desirable." (Rorty 1989, xvi)

What primarily moves us now, he contends, are not theoretical essays or logical proofs, but compelling works of art – novels, movies, plays, and TV programs. (Rorty 1989, xvi) Absent of any privileged 'meta-vocabularies' or objectivity, there is only the constant power struggle to move and to motivate – and in this, fictional appeals to the imagination are often more powerful than factual appeals to reason.

But if there is no ultimately good reason to be a liberal (or conversely, *not* to be a Nazi, Stalinist, fascist, or theocrat) then the state teaching of morals may appear perverse. To promote, sometimes coercively, the values of racial and gender equality, secularism, and inalienable rights – even while admitting that none of these are 'true' – will strike many as tyrannical.

Imagine the school assembly presenter who proudly moralizes before their adolescent audience, hands down strict rules of conduct and severe penalties for transgressing said rules; Imagine, then, that this person proudly announces that they "have no ultimately good reasons for what they say." One person's wry, humane irony becomes another's arbitrary yoke.

The contemporary complaint against "liberal fascism" (Goldberg 2009) is wildly anachronistic. However, one may be forgiven for calling Rorty's politics a type of 'liberal chauvinism.' For it explicitly demands that society be organized a certain way, merely because he wills it to be so. This is the consequence of his taking pluralism to its logical end. If there is nothing but a plurality of worldviews, with no objective, rational basis for choosing between them, then all we are left with is the political decision itself. The liberal, then, seems to substantiate the libertarian's skepticism: The moralizing state coerces without justification, and cannot forge value-consensus through common reason.

Certainly, not all pluralists fall into this secular, liberal tradition. The emphasis on decision and imagination (rather than universal reason) can very well invert itself. This is the conceptual basis for the other major tendency within pluralism – that of post-secularism. Exploiting the central tensions of liberal pluralism, the post-secularist will often point to the conflict between concrete, (often religious) lived experiences, on the one hand, and secular 'reasonableness' on the other.

The latter, they charge, falsely presents itself as objective and universal, superseding the "integrity of different facets of life," especially faith-based ones. Official secularism tends towards homogenization in "a single principle" so that the "door is barred against further discovery." (Taylor 2007, 771–72, 769) Charles Taylor, a major proponent of post-secularism, goes so far as to describe the secular landscape as a sterile waste land. The irrepressible human drive toward the sacred will mean that, sooner or later, "young people will begin again to explore beyond the boundaries" of this restrictive environment. (Taylor 2007, 770)

The post-secularist sees liberal politics, itself, as implicitly relying on sacred vocabularies for its continued relevance. Jürgen Habermas gives clear expression to this view.

> … I do not believe that we, as Europeans, can seriously understand concepts like morality and ethical life, persons and individuality, or freedom and emancipation, without appropriating the substance of the Judeo-Christian understanding of history in terms of salvation. And these concepts are, perhaps, nearer to our hearts than the conceptual resources of Platonic thought, centering on order and revolving around the cathartic intuition of ideas...But without the transmission through philosophy of *any one* of the great world religions, this semantic potential could one day become inaccessible. (Habermas 1992, 15)

Religious vocabularies are not incidental, but essential, for politics. They speak to the imagination and the heart, without which the discursive norms of liberalism become distant, impotent abstractions. Rorty's militant decision in favor of the secular is reversed by his own premises. If what matters is the emotionally compelling, rather than the objectively convincing, what could be more powerful than fervent, religious faith?

Still, one should not overdo the supposed differences between these two pluralist tendencies. For his part, Étienne Balibar is keen to give liberal universalism its due, lest the criticism against it become, itself, too abstract and ahistorical. (Balibar 2018, 14) Meanwhile, Charles Taylor explicitly accepts the Rawlsian notion of overlapping consensus when it comes to international human rights. As he puts it, "we would agree on the norms while disagreeing on why they were the right norms, and we would be content to live in this consensus, undisturbed by the differences of profound underlying belief." (Taylor 1999, 101)

The right to education, housing, and healthcare is less controversial if untied from theory-laden, supposedly Western concepts, of 'human dignity' and 'universal emancipation.' Wide acceptance of such rights across diverse cultures is more likely, contends Taylor, if "the [philosophical] package could be untied" from the legal norm, so that one "could be adopted alone, without the other." (Taylor 1999, 105)

As always, the pragmatic presentation conceals a more substantive outlook. The post-secularist not only wants to triangulate agreement across cultures, but really sees diverse cultures and religions as no less universal than liberal secularism itself. Balibar, especially, sees the tradition of state secularism in France (*laïcité*) as tending toward its own particular, sacralized language, especially when it finds itself opposing traditional religious forms. (Balibar 2018, 20) 'The Republic' has come to replace Mother Church as the locus of impassioned belief and anxious defense against dangerous nonbelievers. One need look no further than the restrictions against religious clothing in public spaces in France (and the Islamophobic rhetoric surrounding those measures). French, secular republicanism, says Balibar, often manifests itself as a national, civic religion.

(Balibar 2018, 120) As such, it becomes rather arbitrary to call religions "particular" and secularism "universal." (Balibar 2018, 22)

Balibar emphasizes the suffix in the word cosmo*political*, as the modern cosmopolis is marked by political negotiation, critique, and agonism as much as an inclusive universalism. (Balibar 2018, 20) The post-secularist thus ends up in a similar position as their liberal counterparts. Religious differences are seen as irreconcilable "differends" which cannot be bridged or wholly criticized from a rationalist perspective. At most, the state can act as a "vanishing mediator" between religious identities, offering immanent criticisms and pointing out tensions, but never imposing its own comprehensive doctrine from without. (Balibar 2018, 53–54, 60) In a spirit of humility, we must seek a "secularization of the secular." (Balibar 2018, 50)

Pluralists of all kinds accept a morality-infused politics. They say "yes" to question 1. This is what separates them from the libertarian. They admit that statecraft – making laws, crafting school curricula, and setting budget priorities – necessarily involves a series of value judgments. But, at the same time, pluralists consider metaphysical certainty (i.e., certainty about the nature of reality) to be a dangerous illusion. For Rorty, as we have seen, this implies a stark decision (without justification) in favor of one, particular value system, namely, secular liberalism. For post-secularists, like Balibar, it means a conception of civility based on the play of multiple cultural identities. (Balibar 2018, 158)

But what the most consistent pluralists share in common is a denial that their politics can be ultimately justified – that their values have any objective grounding in reality as such, outside of a particular time, place, or tradition. Justification instead gives way to imagination, faith, and the political imaginary of the now. Whereas the libertarian took public reasoning out of morals, the pluralist takes Reason (if not humble 'reasons') out of public deliberation. But for all that, what remains is the same intractable contest of wills.

The Rigorous Republican ("Yes" to all three questions)

Suppose, then, that we cut to the chase. If we want something more than arbitrary decisionism, or the contest of imaginations and wills, then we need to get back to the Truth with a capital 'T.' We need an objective, governing philosophy to ground our secular politics. What's more, it is incumbent upon the state to instill in citizens, even through coercion, the right sort of beliefs. This is the tradition of "rigorous republicanism." (Moggach 2003a) For the purposes of our schema, the rigorous republican is one who answers "yes" to all three of our questions: (1) The state must promote certain values over others; (2) It must have ultimate justifications for said values; and (3) The state should compel its citizens to believe in these ultimate justifications.

The rigorous republican, true to their name, embodies a kind of enthusiastic consistency. If the state is in the business of teaching ethics, then it ought to have good reasons for what it teaches and enforces. If these good reasons exist, then

lawful citizens should be made (in whatever ways are most effective) to embrace them. Otherwise, we are asking people to follow laws blindly, or merely to pay insincere lip-service to what is right and just.

Of course, the danger of such a view is easy to detect. Rigorous republicanism demands, not merely obedience to the law, but also a form of 'confession' from private citizens. Unapologetically in the business of ethics, the state claims sovereignty over people's outer actions *as well as their inner beliefs.* Both are equally subject to official control and censure. Rigorous republicanism, therefore, has the tendency to become despotic. As the infamous prosecutor of the Moscow show trials, Andrey Vyshinsky, boasted: "We draw no distinction between intention and the crime itself. This is an instance of the superiority of Soviet legislation to bourgeois legislation." (Solzhenitsyn 1973, 72)

The rigorous republican position, while severe, has a long history. It was the ideology of one of the most radical parties of the French Revolution, the Hébertists, often styled as 'the exaggerators.' Like the also revolutionary (and better known) Jacobins, the Hébertists stood for the abolition of the monarchy, the revocation of Church privileges and property, and the displacement of what they saw as Christian superstition in favor of secular virtue. However, they went beyond the Jacobins in their level of violent hostility towards the Catholic Church. Hébertists engaged in the routine killing of priests, and even banned priestly dress in public spaces. (Fluss 2016) The point was not merely to revoke official church privileges, but also to coerce people out of their privately held beliefs. All of this indicates the potential of a supposedly 'radical' secularism to turn illiberal, authoritarian, and intolerant.

But what happens when the target of rigorous republicanism is not a hegemonic institution like the Catholic Church of 18th Century France, but instead, an oppressed minority? In the 19th Century, the philosopher Bruno Bauer famously argued against political emancipation for Germany's Jews. He did so, not for traditionalist reasons of wanting to convert the Jews to the official state religion of Lutheranism, but instead, upon a strictly secular basis. A member of the 'Young Hegelians,' Bauer saw the republican state as the embodiment of freedom itself. He opposed all special privileges for individual groups, and for this reason, could not abide either the Christian identity of the confessional state, nor the special emancipation of the Jews 'as Jews' (i.e., as a particular group).

What Bauer wanted was a thoroughly secular state which embodied pure freedom itself. By its nature, the free state could not recognize this or that group or interest. As against all historical forms of hierarchy and tradition, Bauer's republicanism stood for the limitless critique of everything according to a pristine, unbounded sense of reason. (Moggach 2022) Accordingly, the merely Christian state has no business emancipating anyone (because it is not yet, itself, free). But equally, the Jews (if they desire to remain a "people apart," with their particular religious laws and customs) are unfit to be emancipated. (Bauer 1983b, 188, 196–97)

In any case, even if the Jews affirm their new status as secular citizens, this will only be an expedient lie for the sake of increased civil rights. Bauer claims of the emancipated Jew, "...he is and remains a Jew in spite of his being a citizen. For his narrow Jewish nature always in the end triumphs over his human and political obligations." (Marx 1977, 41 cit. Bruno Bauer, "The Capacity of Present-Day Jews and Christians to Become Free") What is required, therefore, is the total, social emancipation of the whole people. This goes beyond 'equality before the law,' and entails a full reformation of the soul. We are all to become secularized through and through before deserving equal political rights.

After this utopian vision failed to materialize in the 1848 Revolutions, Bauer turned sharply to the Right. If democracy, liberalism, and individual rights could not bring about the rule of universal reason, then what is needed is the unitary rule of a world-imperialist power. Bauer predicted the coming of a cataclysmic world war which would bring about a transnational sort of imperialism. (Moggach 2022) It is within this brave new world that society would become massified and homogenized, and only out of this political "dust" could a new, fully liberated elite emerge. Free of the traditions and constraints associated with the "old world," this elite could finally engage in unbridled acts of creation and self-consciousness. But of course, this ecstatic "freedom" will be at the expense of those ordinary people who fail to become elite themselves. (Bauer 1983a)

If it wasn't always clear at the time that Bauer's rigorous republicanism accommodated reactionary politics, it became abundantly obvious in the 20th Century. The Nazi jurist Carl Schmitt embraced Bauer's critique of traditional, conservative values; For these are merely passive. The slavish devotion to established tradition is "an irrational rejection of every intellectually conscious decision." (Schmitt 2005, 54) What counts, for Schmitt as well as Bauer, is rather the strong individual's free, self-creation. The sort of 'Reason' which can found a state is so pure and singular that it is evacuated of all determinate content or meaning. It loses its objective character and becomes an immaculate act of the will.

Of course, not everyone can be so free. Schmitt affirmed the Catholic doctrine of papal infallibility, but now as a secularized political concept to be applied to the state leader. Accordingly, Schmitt's "leadership principle" (*Führerprinzip*) meant that the sovereign cannot be constrained by either tradition or law. The masses, for their part, must be thoroughly depoliticized and conform to the unbounded dictates of the *Führer*. (Gross 2007, 23, 52, 102, 167, 175–76) We have here the modern confluence of the unitary state, unbounded by tradition, and unified by the free will of a singular leader. The people express themselves politically (if at all) through their wholehearted embrace of this leader and the values he spontaneously creates. The upshot is the same as in Bauer's 19th Century politics: a sincere unity of spirit, not mere conformity to outer law.

All of this may seem like boilerplate fascism – and it is. But these ideas are also informed by Schmitt's critique of modern political philosophy, especially his reading of the 17th Century English philosopher Thomas Hobbes. For it was

Hobbes who affirmed the supremacy of the prince to dictate matters of law as well as belief. (Schmitt 2008, 56) Yet, in a moment of high-minded generosity, Hobbes allows commoners the right to pay 'lip service' to the religious claims of the sovereign. Can the king (as tradition states) truly cure diseases through the laying on of hands? Who knows? All that matters is that the commoners act *as if* they believed; it's simply a matter of public respect.

But it is precisely this modicum of toleration, this tiny space left for dissenting thought, that Schmitt identifies as self-destructive of the state. He traces the problem, in particular, to Baruch Spinoza, that "alien" Jewish philosopher who exploited the small gap left by Hobbes between outer piety and inner belief. Through this gap, Spinoza develops a whole litany of individual, liberal rights (freedom of thought, i.e. "the freedom to philosophize," first among them). It is this separation, naively opened by Hobbes, but cunningly exploited by "the Jew" Spinoza, that corrodes the State from the inside. (Schmitt 2008, 57)

In Schmitt's Nazi ideology, this treachery is not limited to Spinoza alone, but rather is the *modus operandi* of all modern, secular Jews. Hence the problem with liberal toleration for unassimilated minorities; They will always take advantage at the expense of the body politic. Aping Bauer, Schmitt insists the Jew may act like a real citizen, but privately remains a self-interested member of his own clan. (Gross 2007, 160) The Medieval blood libel is recast in political terms: The Jew sacrifices and divides the state in order to safeguard his own special interests. (Schmitt 2008, 90) The state is emasculated by excising 'belief' as its vital, animating force. And the Jews, who can never be more than internal aliens, destroy the unity and coherence of political life through modern doctrines of liberalism (especially the freedoms of thought, expression, and religion).

In this way, rigorous republicanism may start from a position of pure freedom, but that's not where it ends up. For the state (that locus of freedom on earth) can only function if private citizens actually affirm its fundamental principles 'in their soul,' as it were. After all, action is motivated by thought. Therefore, outer loyalty can only be sustained through a conformity of inner belief. Mere 'lip service' isn't enough, as a true confession must be sincere and heartfelt. Dostoevsky's character from *Demons*, Shigalyev, expresses the paradox of rigorous republicanism best: "…my conclusion directly contradicts the original idea I start from. Starting from unlimited freedom, I conclude with unlimited despotism." (Dostoevsky 1995, xix)

The Rationalist Republican ("Yes" to questions 1 and 2; "No" to question 3)

If 'rigorous republicanism' turns despotic, this leaves us with only one remaining permutation: One must say "yes" to questions 1 and 2, but "no" to question 3.[5] In

[5] Recall that each question in the series is logically a 'gate' for those further down the list, where answering "yes" to question 1 is the precondition for even considering question 2, etc. That is

other words, the state will promote certain ethical norms over others; it should have good, ultimate reasons for doing so; yet crucially, it must *not* compel its citizens to personally endorse those ultimate reasons. This is the nuanced position we will call 'rationalist republicanism.' The state teaches ethics but does not demand forced confession.

The rationalist republican earns their name by affirming a secular 'state philosophy.' This is a rationally justified worldview which motivates its various laws, policies, and priorities. Such a position will be jarring for many, and it evokes in the liberal imagination visions of thought-police, totalitarian control, and all manner of Orwellian nightmares. However, unless public laws are based on some sort of rational justification, these will be unmoored and arbitrary – precisely the deficit of the 'pluralist' position examined above.

Moreover, unless these ultimate justifications are codified, transparent, and available to the general public, then such a 'state philosophy' will be reduced to an esoteric doctrine of the political elite. Laws need a rationale, and all people subject to the law deserve to know what those supposedly good, foundational reasons are. Justification and education are thus indispensable as they are linked. Both the letter and the animating spirit of the law have to be explicable to the general public if participatory democracy is to have genuine meaning. The average citizen should know, not only *what* the law says, but also *why*.

At the same time, the rationalist republican draws a sharp line between a normative 'civic education,' on the one hand, and forced 'confession' on the other. The former is about explicating the core values of the state. In other words, what view of reality, human nature, and the good motivates our norms and policies? Forced confession is another matter entirely. This involves compelling individuals to publicly assent to this worldview. The rationalist republican eschews such coercion and maintains, as sacrosanct, the freedom of conscience.

This balanced position was first clearly articulated by the 17th Century philosopher, Baruch Spinoza. In his major political work, the *Tractatus Theologico-Politicus* (TTP), Spinoza outlines a potent vision of the state alongside a strict tolerance for individual, non-conforming belief. For example, he argues that the state must have undivided authority over its citizens, and not share its right to police public behavior with any private group. (Baruch Spinoza 2007, 193)[6] Spinoza had in mind, in particular, the various religious sects of his day which sought to exercise their own authority over private citizens, and so undermine state power. (Baruch Spinoza 2007, xxvi–xxvii)

This need to maintain state sovereignty, over and above private, religious interests, led Spinoza to adopt a form of 'state religion.' This is the doctrine known as Erastianism, an early modern theory that only the state, and not private

why the rationalist republican position ("yes" to 1 and 2; "no" to 3) is the only remaining combination of responses.

[6] All references to the main body of Spinoza's *Tractatus Theologico-Politicus* use the Gebhardt page numbers.

churches, should have the right to punish sins.[7] But far from being theocratic, Spinoza's vision was one of secular authority. Only the state can determine the proper ways citizens should behave toward one another. Interpersonal actions (what he calls "outer piety") are a matter of public law and not private, religious censure. This means that the state must assert its own, comprehensive doctrine of the good which it promotes through civil laws and penalties. It is therefore "the duty of the sovereign authority alone to lay down how a person should behave with piety towards his neighbor, that is, how one is obliged to obey God." (Baruch Spinoza 2007, 232)

Beyond this, the 'state cult' imposes no other articles of belief. As Spinoza puts it, "...the safest policy is to regard piety and the practice of religion as a question of works alone, that is, as simply the practice of charity and justice, and to leave everyone to his own free judgment about everything else." (Baruch Spinoza 2007, 226) Still, this minimalism can be misleading. Spinoza is no pragmatist and does not reduce the public good to 'whatever works.' His politics are entirely grounded by a rationalist philosophy, namely the proposition that reason is "deduced from the very essence of man," is his "greatest good," and is also "common to all." (Benedictus de Spinoza 1985, E4P36 Schol.)[8] The positive mission of the state is thus to allow for our common reason to flourish. Still, the state's coercive powers only extend to the *outer expressions* of this rationalism, namely, acts of solidarity, charity, and justice.

This, then, is the hallmark of the rationalist republican position. It does not pretend that the state can, in libertarian fashion, remain value-neutral; and contra the pluralist, it maintains that there is a comprehensive philosophy which undergirds the law. But at the same time, unlike the rigorous republican, private belief (or 'inner piety') remains beyond the reach of state control. The state affirms a doctrine of the good, and even engages in public indoctrination. But whereas the state does indeed 'teach ethics,' it may never coerce inner belief or demand personal confession.

Spinoza here anticipates the position of Maximilien Robespierre, leader of the radical Jacobin faction during the French Revolution. Despite Robespierre's often blood-soaked reputation as a fanatic, his actual stance on individual liberties was far more nuanced. He opposed, for example, the Hébertist policy of attacking priests for wearing religious garb. And while Robespierre did support a state religion to displace the hegemonic (and counter-revolutionary) Catholic Church,

[7] The modern usage of 'Erastianism' as denoting a subservience of the church to the secular state differs from the original intent of its namesake, Erastus (orig. Thomas Lieber). The latter was, himself, a devout follower of Zwingli's Reformation theology and not, in any meaningful sense, a secularist. His concern had more to do with the narrower question of 'excommunication' within the church.

[8] All references to Spinoza's *Ethics* follow the standard convention of indicating the work itself with the letter "E," propositions with "P," definitions with "D," and axioms with "A," each followed by their respective number.

this was likewise minimalist in nature. His "Cult of the Supreme Being" was, in true Spinozist form, primarily concerned with public acts, or "outer piety," rather than inner belief. The 1794 decree establishing the cult speaks of "the practice of the duties of man" as the best service to the Supreme Being. These are, first and foremost: "to detest bad faith and despotism, to punish tyrants and traitors, to assist the unfortunate, to respect the weak, to defend the oppressed, to do all the good one can to one's neighbor, and to behave with justice towards all men." (Fluss 2016)

Limiting state authority to such outer acts was only prudent. In Spinoza's naturalistic view, political "right" extends only as far as actual "power" allows. Big fish have the "right" to eat little ones, and so forth. (Baruch Spinoza 2007, 189, 240) But as the state hasn't the power to control one's inner thoughts, it therefore has no right. And moreover, it is foolish (and dangerous) to even try. For setting up laws which cannot be reliably enforced only serves to weaken the state and make it look foolish (images of the 'emperor with no clothes' come to mind). As Spinoza puts it, "Trying to control everything by laws will encourage vices rather than correcting them. Things which cannot be prevented must necessarily be allowed, even though they are often harmful." (Baruch Spinoza 2007, 243) And while interpersonal communication is, strictly speaking, an outer act, it is so close to inner thought that this too should be generally free from state control. As such, the "liberty to philosophize" also includes the liberty to philosophize *out loud*. (Baruch Spinoza 2007, 243)

In Spinoza's view, even if speech could be strictly controlled, doing so would undermine the stability of the state. In such a case, people's compelled speech would be constantly divorced from their true, inner thoughts. Citizens "would be continually thinking one thing and saying something else…" And this, in turn, would lead to a flourishing of deception and flattery, and undermine precisely the sort of trust needed to maintain a functioning society. (Baruch Spinoza 2007, 243–44)

In any case, laws which censor speech most often affect the wrong sorts of people. Unscrupulous self-seekers will say whatever is expedient in order to get ahead, but only the virtuous, free-minded person will be willing to stand on principle despite dire consequences. This can only show the sovereign state in a bad light.

> What greater ill can be devised for any commonwealth than for honest men to be banished like outlaws because they think differently from the rest and do not know how to hide this? What is more dangerous, I contend, than for people to be treated as enemies and led off to death, not for misdeeds or wrongdoing, but because they make a free use of their intelligence, and for the scaffold which should be the terror only of wrongdoers to become a magnificent stage on which to exhibit to all a supreme exemplum of constancy and virtue while casting the deepest reproach to the sovereign? (Baruch Spinoza 2007, 245)

Nearly a century later, in revolutionary France, Robespierre would strike a similar chord. He likewise maintained that it was imprudent for the state to try and coerce people out of their traditional, religious beliefs. Even if such beliefs are sometimes fanatical, anti-social, or hostile to the common good, repressing them will only breed resentment and obstinacy. It is a fact that much of the Catholic Church in 18th Century France (especially its upper echelons) actively sought to undermine the Revolution and reinstate the *Ancien Régime*. Yet, for Robespierre, this did not warrant the abolition of personal Catholic beliefs or praxis. In his words, "I see only one way of reviving fanaticism among us: it is to affect to believe in its power…Priests have been denounced for saying the Mass. They will continue to do so all the longer if you try to prevent them. He who wants to prevent them is more fanatical than the priest himself." (Rudé 1976, 123)

This line of thinking extends to the 19th Century writings of Karl Marx and Friedrich Engels. Their communism, likewise, issued from a secular vision of the world known as 'historical materialism.' Such materialism is incompatible with traditional religious thought, and each in turn implies a different political outlook. All the same, Engels has only bitter polemics for those utopian-minded socialists who wish to 'abolish' religion outright. For this only "helps it [religion] to martyrdom and a prolonged lease of life." (Engels 2018, 355) As he sarcastically asserts elsewhere, "...the only service, which may still be rendered to God today, is that of declaring atheism an article of faith to be enforced and of outdoing even Bismarck's anti-Catholic laws by forbidding religion altogether." (Engels 2002)

This debate over religious toleration came to a head over the so-called 'Jewish question.' Whereas Bruno Bauer, as we saw, demanded that the Jews give up their religion as a prerequisite for political emancipation, Karl Marx would have none of it. Here, the two versions of republicanism (rigorous and rationalist) confronted one another on German soil. In his famous polemic against Bauer, *On the Jewish Question* (1843), Marx insists that the equal rights of political emancipation come before (not after) a critique of religious belief. "Bauer asks the Jews: Does your standpoint give you the right to seek political emancipation? But we ask the reverse question: Has the standpoint of political emancipation the right to require from the Jews the abolition of Judaism and from all men the abolition of religion?" (Marx 1977, 42–43) In other words, for Marx, equal political rights do not require a full transformation of the soul. To the contrary, civil rights are merely the first, minimal step in establishing the sort of society where – one day – a genuinely social (i.e. 'human') emancipation might be possible.

Underlying this critique is the more fundamental point that social relations and material circumstances pattern individual belief. Contrary to Bauer's ecstatic notion of reason, the intellect is not a spontaneously free self-creation. It is, rather,

entirely bound up with material conditions. And it's this insight which makes all the difference when it comes to political toleration.[9]

The rigorous republican (since unlinking reason from matter), can only see irrational beliefs as something like a willful sin for which we are morally culpable. For nothing determines the mind apart from itself, i.e., its own free will. No wonder, therefore, that Bauer insists that the state should demand fealty, not only in outer action, but also in inner belief. Anything less would be tantamount to treason. But the rationalist republican (from Spinoza through Marx and Engels) recognizes that false beliefs are not so much 'chosen' as they are 'caused.' Thus, whereas Engels has no admiration for traditional religion, neither is he interested in chastising individual believers.

Better, instead, to diagnose the material and social causes of supernatural belief. Religion exists, on this Marxist account, because we still confront a society and economy in which we are alienated – alienated from one another, from the means of production, and from the very products of our own labor. (Marx and Engels 1988, 74–75) The market renders us cogs in a machine where we work for the capitalist's profits rather than for human needs, and so we compete with our fellow workers for scarce resources and employment rather than cooperate for the common good. This is how our worldly alienation finds its mirror-image in mythos. Our collective imagination translates powerlessness in the material world to powerlessness before an omnipotent Father. As Engels puts it, "All religion…is nothing but the fantastic reflection in men's minds of those external forces which control their daily life, a reflection in which the terrestrial forces assume the form of supernatural forces…" (Engels 2018, 353)

Logically, then, irrational beliefs are most effectively combatted through a "social act" rather than direct, state censorship. (Engels 2018, 355) It is the democratic taking control of the economy (the forces of production and distribution) that gradually removes the compulsion to imagine some transcendent force above us, controlling us. Religion, for Marx and Engels, is not abolished but instead withers away. Here, Engels is quite clear about the order of operations:

> And when… society, by taking possession of all means of production and using them on a planned basis, has freed itself and all its members from the bondage in which they are now held by these means of production which they themselves have produced but which confront them as an irresistible alien force… only then will the last alien force which is still reflected in religion vanish; and with it will also vanish the religious reflection itself, for the simple reason that then there will be nothing left to reflect. (Engels 2018, 355)

[9] It is crucial to point out here that Bruno Bauer's anti-socialism and his anti-Semitism issue from the same place. Both socialists and the Jews, he complains, are preoccupied with material concerns. By contrast, genuine liberation involves a constant overcoming of all objective, material conditions. (Moggach 2003b, 444–45)

Similarly, for Marx, full human emancipation means the recognition of our own powers, and the denial of any transcendent power above us. This is fundamentally bound up with asserting our common, social control over our material circumstances. Thus, "man must recognize his own forces as social forces, [and] organize them… Only when this has been achieved will human emancipation be completed." (Marx 1977, 57) Once more, the order of things is of paramount importance:

> We no longer regard religion as the cause, but only as the manifestation of secular narrowness…We do not assert that they [individuals] must overcome their religious narrowness in order to get rid of their secular restrictions, we assert that they will overcome their religious narrowness once they get rid of their secular restrictions. We do not turn secular questions into theological ones. (Marx 2009)

This 'rationalist republican' character of Marxism persisted through the 20th Century, notably in the works of the Russian revolutionary Leon Trotsky.[10] In *Results and Prospects*, he inveighs against the 'shallow moralizing' of socialist ideologues. These moralists, in the model of Bruno Bauer, see the creation of a new sort of human being – virtuous and free of egotism – as the precondition for socialist revolution. To this frame of mind, Trotsky objects that the real order of things is exactly the reverse: "Socialism does not aim at creating a socialist psychology as a pre-requisite to socialism but at creating socialist conditions of life as a pre-requisite to socialist psychology." (Trotsky 2010, 108-9) In other words, we don't need to become high-minded angels to end exploitation; Rather, it is the end of capitalist exploitation that allows humanity to gradually emerge from its self-imposed egotism and selfishness.

None of this implies that, for the rationalist republican, politics is divorced from belief. Ideas matter. The state must maintain some 'comprehensive doctrine' of reality and the good. For this is the justification and rationale for public policy. And, moreover, these doctrines can and should be widely disseminated, whether in the form of a 'public cult' or merely within school curricula and courses on civic education. The state *does* teach ethics. But because inner beliefs are often conflicted, fragmentary, and conditioned by material circumstances – and because they are, in any case, beyond the actual control of the law – then acceptance of the 'state philosophy' can never be a matter of coercion, punishment, or forced confession. Freedom of conscience remains intact, even though the state ceases to pretend to be value-neutral or philosophically agnostic.

[10] Some may object to characterizing Marxism as a species of 'republican' thought. After all, Marxism advocates the withering away of the state, not its supremacy over private interests or ideologies. However, the term 'state,' used in this way, is always meant narrowly, i.e., as "an organ for the oppression of one class by another." (Lenin 1992, 9) But if, as in this article, we define the state as simply an "official, public administration," then this appears indispensable to the Marxist project of organizing collective control over the means of production.

But there is, finally, a deeper reason why the rationalist republican wishes to safeguard the 'freedom to philosophize.' This has to do with the nature and purpose of the state itself. Recall that, for Spinoza, reason is the very thing which defines humanity. (Benedictus de Spinoza 1985, E4P36 Schol.) It is that singular quality which allows humans to work together, increase their wellbeing, and flourish. (Benedictus de Spinoza 1985, E2P38 Cor., E4 Appendix 9) The state, therefore, has a very specific mission, namely, to facilitate this rational faculty among its citizens. As Spinoza puts it, "the true purpose of the state is in fact freedom." (Baruch Spinoza 2007, 241) But for him, freedom is wholly inseparable from reason. It is the ability to order our own thoughts and actions, rather than being constantly swayed by external stimuli and passions.

This being the case, freedom of conscience is more than just prudent statecraft. Policing thought is destructive of the state's highest function as the guardian and promoter of human reason. This is because the development of one's rational faculties is not the sort of thing that can be directly forced. It requires, instead, the ability to experiment, reflect, revise, and make mistakes without constant fear of external coercion or censorship. As Spinoza states, "this liberty is absolutely essential to the advancement of the arts and sciences; for they can be cultivated with success only by those with a free and unfettered judgment." (Baruch Spinoza 2007, 243) A state which does not allow reason to improve organically, which does not grant the negative freedom to form erroneous beliefs and false ideas, thereby undermines its very purpose. It blunts, not only the various products of reason (the arts and sciences), but the very *capacity* to reason itself.

Conclusion

Looking back at the intellectual terrain sketched above, one can discern the unique virtues of the rationalist republican position. In contrast to the libertarian and pluralist, the rationalist republican maintains that the state must both (1) promote certain ethical values over others, and (2) have some ultimate justification for those values it promotes. The law is based on a specific philosophy or worldview. Yet, in contrast to its 'rigorous' counterpart, the rationalist republican answers "no" to question 3: Citizens should not be personally compelled to accept the state's philosophy as against their own conscience. As such, only the rationalist republican avoids the twin pitfalls of skepticism and despotism. The laws are based on well-founded reasons, even while individual citizens need not confess their belief in said reasons.

Furthermore, only the rationalist republican provides a sure basis for political toleration. While the rigorous republican is openly hostile to dissenting beliefs, the pluralist proves to be an 'unreliable friend' of the dissenting minority. In eschewing a positive, publicly intelligible conception of the good, the latter can only argue from negative premises: One tolerates dissent because one is never sure about the actual truth. Rawls' 'burdens of judgment' here function in the same

manner as Balibar's focus on irresolvable cultural differences. And these are functionally similar to Rorty's outright denial of rational justification. In each case, a humane, tolerant politics is thought to be built upon the negative claim that, when it comes to objective truth, "we simply can't know for sure."

But one cannot draw positive conclusions from such negative premises. A lack of objective knowledge – about nature, the good, and human nature – need not imply a tolerant politics. It more readily implies the opposite. In the absence of universal truths, what is left is nothing but the contest of free, creative wills who invent, rather than discover, what is 'true.' In that case, it is the strong will that prevails, entirely unconstrained by notions of the common good, equality, or toleration. For there are no objective criteria by which to judge the free, political decision.

By contrast, it is the rationalist republican that provides a positive basis for political toleration. In identifying reason as universal and essential to all individuals, it confers a specific mission to the state: to extend and empower human reason to the greatest extent possible. That is why political dissent must be vigorously defended – not because we don't know the truth – but rather because we know that human beings are essentially rational, and that the flourishing of reason requires the liberty to sometimes get things wrong.

References

Balibar, Étienne. 2018. *Secularism and Cosmopolitanism: Critical Hypotheses on Religion and Politics*. European Perspectives. New York: Columbia University Press.

Barnett, Randy E. 2004. "The Moral Foundations of Modern Libertarianism." In *Varieties of Conservatism in America*, edited by Peter Berkowitz, 51–74. Stanford, CA: Hoover Institution Press.

Bauer, Bruno. 1983a. "The Genus and the Crowd." In *The Young Hegelians: An Anthology*, edited by Lawrence S. Stepelevich, 198–206. Cambridge, UK: Cambridge University Press.

————. 1983b. "The Jewish Problem." In *The Young Hegelians: An Anthology*, edited by Lawrence S. Stepelevich, 187–97. Cambridge, UK: Cambridge University Press.

Berlin, Isaiah. 2013. *The Crooked Timber of Humanity: Chapters in the History of Ideas*. Edited by Henry Hardy. Second Edition. Princeton, NJ and Oxford: Princeton University Press.

"Bullying Prevention Training Center." n.d. Stopbullying.Gov. Accessed February 14, 2022. https://www.stopbullying.gov/resources/training-center.

Cherniss, Joshua, and Henry Hardy. 2022. "Isaiah Berlin." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. URL = <https://plato.stanford.edu/archives/spr2022/entries/berlin/>.

Dostoevsky, Fyodor. 1995. *Demons*. Translated by Richard Pevear and Larissa Volokhonsky. New York: Vintage Books.

Engels, Frederick. 2002. "The Program of the Blanquist Fugitives from the Paris Commune." Translated by Ernest Untermann. Marxists Internet Archive. 2002. https://www.marxists.org/archive/marx/works/1874/06/26.htm.

————. 2018. *Herr Eugen Dühring's Revolution in Science: Anti-Dühring*. Translated by Emile Burns and C.P. Dutt. Classic Reprint. London: Forgotten Books.

Fluss, Harrison. 2016. "Revisiting the Cult of the Supreme Being." *Jacobin*, January 14, 2016. https://jacobinmag.com/2016/01/robespierre-rousseau-religion-separation-church-state-kim-davis.

Goldberg, Jonah. 2009. *Liberal Fascism: The Secret History of the American Left from Mussolini to the Politics of Change*. New York: Broadway Books.

Gould, Stephen Jay. 1999. *Rocks of Ages: Science and Religion in the Fullness of Life*. New York: Ballantine Books.

Gross, Raphael. 2007. *Carl Schmitt and the Jews: The "Jewish Question," the Holocaust, and German Legal Theory*. Madison, WI: University of Wisconsin Press.

Habermas, Jürgen. 1992. "Metaphysics after Kant." In *Postmetaphysical Thinking: Philosophical Essays*, translated by William Mark Hohengarten, 10–27. Cambridge, MA; London: MIT Press.

Kelley, David. 2001. *The Contested Legacy of Ayn Rand: Truth and Toleration in Objectivism*. New Brunswick, NJ: Transaction Publishers.

Kierkegaard, Søren. 1997. *Kierkegaard's Writings, IX, Volume 9: Prefaces: Writing Sampler*. Edited and translated by Todd W. Nichol. Princeton, NJ: Princeton University Press.

Land, Nick. 2012. "The Dark Enlightenment." *The Dark Enlightenment* (blog). December 25, 2012. http://www.thedarkenlightenment.com/the-dark-enlightenment-by-nick-land/.

Lenin, V.I. 1992. *The State and Revolution*. Translated by Robert Service. London; New York: Penguin Books.

Marx, Karl. 1977. "On the Jewish Question." In *Karl Marx: Selected Writings*, edited by David McLellan, 39–62. Oxford: Oxford University Press.

————. 2009. "On The Jewish Question (1843)." Edited by Andy Blunden, Matthew Grant, and Matthew Carmody. Marxists Internet Archive. https://www.marxists.org/archive/marx/works/1844/jewish-question/.

Marx, Karl, and Frederick Engels. 1988. *Economic and Philosophical Manuscripts of 1844 and The Communist Manifesto*. Translated by Martin Milligan. Amherst, NY: Prometheus Books.

Meyers, Lawrence. 2011. "Politics Really Is Downstream from Culture." *Breitbart*, August 22, 2011. https://www.breitbart.com/entertainment/2011/08/22/politics-really-is-downstream-from-culture/.

Moggach, Douglas. 2003a. "Republican Rigorism: Hegelian Views of Emancipation in 1848." *The European Legacy* 8 (4): 441–57.

————. 2003b. *The Philosophy and Politics of Bruno Bauer*. Cambridge, UK: Cambridge University Press.

————. 2022. "Bruno Bauer." In *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. forthcoming URL = <https://plato.stanford.edu/archives/spr2022/entries/bauer/>.

Newport, Frank. 2013. "In U.S., 87% Approve of Black-White Marriage, vs. 4% in 1958." Princeton, NJ: Gallup. https://news.gallup.com/poll/163697/approve-marriage-blacks-whites.aspx.

Pal, Alasdair. 2021. "Taliban Replaces Women's Ministry with Ministry of Virtue and Vice." *Reuters*, September 17, 2021. https://www.reuters.com/world/asia-pacific/taliban-replaces-womens-ministry-with-ministry-virtue-vice-2021-09-17/.

Paul, Ron. 2016. "Education System Broken: Let's Try 'Ed-Exit.'" *Ron Paul Institute for Peace and Prosperity* (blog). November 21, 2016. http://www.ronpaulinstitute.org/archives/featured-articles/2016/november/21/education-system-broken-let-s-try-ed-exit/.

————. 2018. "New Tax Plan Makes It Easier to 'Ed-Exit.'" *Ron Paul Institute for Peace and Prosperity* (blog). September 24, 2018. http://www.ronpaulinstitute.org/archives/featured-articles/2018/september/24/new-tax-plan-makes-it-easier-to-ed-exit/.

"Public Schooling Battle Map." n.d. The Cato Institute. Accessed February 14, 2022. https://www.cato.org/education-fight-map.

Rand, Ayn. 1964. *The Virtue of Selfishness: A New Concept of Egoism*. New York: Signet.

————. 1990. "Of Living Death." In *The Voice of Reason: Essays in Objectivist Thought*, edited by Leonard Peikoff, 46–63. New York: Meridian.

Rawls, John. 1999. *Collected Papers*. Edited by Samuel Freeman. Cambridge, MA: Harvard University Press.

————. 2003. *A Theory of Justice*. Revised Edition. Cambridge, MA: Belknap - Harvard University Press.

————.2005. *Political Liberalism*. Expanded Edition. Columbia Classics in Philosophy. New York: Columbia University Press.

Rorty, Richard. 1989. *Contingency, Irony, and Solidarity*. Cambridge, UK: Cambridge University Press.

Rudé, George. 1976. *Robespierre: Portrait of a Revolutionary Democrat*. New York: Viking.

Sanchez, Julian, and David Weigel. 2008. "Who Wrote Ron Paul's Newsletters?" *Reason*, January 16, 2008. https://reason.com/2008/01/16/who-wrote-ron-pauls-newsletter/.

Schmitt, Carl. 2005. *Political Theology: Four Chapters on the Concept of Sovereignty*. Translated by George Schwab. Chicago: University of Chicago Press.

————. 2008. *The Leviathan in the State Theory of Thomas Hobbes: Meaning and Failure of a Political Symbol*. Translated by George Schwab and Erna Hilfstein. Chicago; London: University of Chicago Press.

Solzhenitsyn, Aleksandr. 1973. *The Gulag Archipelago: An Experiment in Literary Investigation*. Translated by Thomas P. Whitney. New York: Harper & Row.

Spinoza, Baruch. 2007. *Theological-Political Treatise*. Edited by Jonathan Israel. Translated by Michael Silverthorne and Jonathan Israel. Cambridge Texts in the History of Political Thought. Cambridge; New York: Cambridge University Press.

Spinoza, Benedictus de. 1985. "Ethics." In *The Collected Works of Spinoza, Volume I*, edited and translated by Edwin Curley. Princeton, NJ: Princeton University Press.

Staff Report. 2021. "Palm Beach County School Board To Ratify Holocaust Education Proclamation." *BocaNewsNow.Com*, October 13, 2021. https://bocanewsnow.com/2021/10/13/palm-beach-county-school-board-to-ratify-holocaust-education-proclamation/.

Taylor, Charles. 1999. "Conditions of an Unforced Consensus on Human Rights." In *The Politics of Human Rights*, edited by Obrad Savić, 101–19. London; New York: Verso.

————. 2007. *A Secular Age*. Cambridge, MA; London: Belknap - Harvard University Press.

Trotsky, Leon. 2010. "Results and Prospects." In *The Permanent Revolution & Results and Prospects*, translated by John G. Wright and Brian Pearce, Revised Edition, 33–136. Seattle: Red Letter Press.

"U.S. Equal Employment Opportunity Commission." n.d. Accessed February 14, 2022. https://www.eeoc.gov/federal-sector/federal-training-outreach.

Walker, Jesse. 2012. "When the Right Helps to Write the Sex Education Curriculum." *Reason*, May 7, 2012. https://reason.com/2012/05/07/when-the-right-helps-to-write-the-sex-ed/.

Weiner, Rachel. 2011. "Ron Paul the Religious." *The Washington Post*, October 14, 2011, sec. Politics. https://www.washingtonpost.com/blogs/the-fix/post/ron-paul-the-religious/2011/10/14/gIQAPdUFkL_blog.html?noredirect=on&utm_term=.3b5533aee01b.

The Social Ontology of Alain Badiou's *Being and Event*

Min Seong Kim

Abstract: The innovation of Alain Badiou's theory of change, which has attracted a great amount of attention from scholars working in disciplines across humanities, social sciences, and art over the past two decades, cannot be appreciated independently of the account of situations prior to an event's irruption, namely, the order of being that is conceived using modern set theory in his treatise on general ontology. Retracing the meticulous systematicity with which pre-evental situations are conceived in *Being and Event*, this paper offers a reconstruction of Badiou's general ontology that points toward the potential therein for articulating an account of structures and situations that may be qualified as *social*.

Keywords: Alain Badiou, contemporary French philosophy, social ontology.

1. Introduction

Being and Event by the French philosopher Alain Badiou is the first volume of what might be called his *Summa Ontologica*, which comprises three books published in France between 1988 and 2018: *Being and Event* (2005), *Logics of Worlds* (2009), and *L'Immanence des vérités* (2018). Few works of European philosophy to have appeared in the last several decades – inclusive, arguably, of Badiou's two sequels to *Being and Event* – match its innovation and ambition, which can be succinctly stated as follows: to do philosophy again in a manner reminiscent of Hegel's *Logic*, by founding it on an account of being qua being that Patrice Maniglier has characterized as "perfectly general, nonspecific theory of what *is*, inasmuch as it *is*, and inasmuch as it is nothing in particular," which is "as valid for nature as it is for culture." (2010, 69)

As dictated by its own ambition, *Being and Event* unfolds systematically from the analysis of pure being. The question motivating Badiou's philosophical endeavor, however, can be stated in more concrete terms. The "fundamental question," Badiou says of his philosophy, "is a very simple one: What exactly is something new? What is novelty? What is creation?" The order in which he proceeded, he explains, is not from an abstract thought of being qua being, but rather from "a living experience of what is something absolutely new and a vivid experience of when something happens. […] And I first experienced this point in my life and only after had to create the concepts to justify and clarify this point." (Badiou and Critchley 2007, 361-62) The account of being to which roughly half of *Being and Event* is dedicated is thus a propaedeutic to addressing matters that truly interests Badiou, that is, the theory of event and its transformative

consequences – conceptualized under the terminology of *truths* – developed in the second half of the book. Badiou does not pretend to be alone in engaging the question of novelty, the question of whether it is possible for there to be an experience of something genuinely new and transformative. Every thought of change, political revolutions, or even artistic creations is intricately tied to that question, such that both those who affirm the possibility of radical novelty and those who deny such a possibility must have at least implicitly engaged with the question – it is certainly "not by chance that French post-structuralist thinkers have put so much weight on the concept of 'event,' both from a theoretical and political perspective." (Renault 2016, 30) But, at the same time, if so many thinkers in the philosophical tradition to which Badiou too belongs have made references to some moment of rupture, and 'event' has been the most common name given to that moment, the designation 'the philosopher of the event' attached to Badiou's name should not occlude the simple truth that whatever that is of value in Badiou's works is to be found in *how* Badiou thinks the event, not *that* he thinks the event.

Thus, while it is understandable that aspects of Badiou's philosophical system that have attracted the great amount of attention from scholars working across humanities, social sciences, and art are those pertaining to novelty, event, and change, the innovation of Badiou's theory of change cannot be appreciated independently of the account of the world before an event, namely, the pre-evental order of being that is presented in *Being and Event* with meticulous systematicity.[1] Indeed, unless an account of structures and situations that could rightly be qualified as social could be reconstructed from Badiou's general ontology, the value of *Being and Event* for conceptualizing novelty and change in the specific situations scrutinized by many disciplines in the humanities, social sciences, and art would be significantly diminished. To show how the seemingly esoteric categories of *Being and Event* can be transposed onto categories relevant for thinking the fundamental structure of the social world, this paper offers a succinct but sufficiently detailed reconstruction of the account of the order of being in *Being and Event* that maintains an emphasis on the feature of that account that has perhaps been the most controversial, as well as constituting an obstacle for many readers whose background is in the humanities: the decision to read mathematics, particularly modern set theory, as a general ontology. It is to be acknowledged that the relation between the strictly set-theoretical ontological discourse and extraontological – including social – discourse is a problematic one in Badiou's philosophy, regarding which some remarks shall be made in the section that

[1] Because Badiou decides that mathematics is the discourse of being, there is no such thing as *Badiou's* ontology in the strict sense. Whenever that expression is employed hereafter, it shall denote the 'metaontological' attempt carried out in *Being and Event* to determine, based on what the ontological discourse – mathematics, or more specifically, set theory – speaks of being, particularly Badiou's philosophical concepts of presentation, representation, void, truth, and subject.

immediately follows. Nevertheless, this paper will have fulfilled its purpose in so far as the connection made between the categories of *Being and Event* and social theory facilitates an understanding of the categories presented therein further than would be possible without referring to its mathematical basis at all, hinting thereby ways in which Badiou's set-theoretical ontology allows for productive theorizations of social situations and social change.

2. Mathematics as the Discourse of Being qua Being

The following inaugural decision of Badiou's ontological discourse is no doubt motivated at least in part by a desire to dispel every vestige of mysticism that the question of being has often tended to evoke: "What can rationally be said of being qua being, of being devoid of any quality or predicate other than the sole fact of being exposed to thought as entity, is said – or rather *written* – as pure mathematics." (Badiou 2004a, 168) While the decision for the identity of the discourses of being and mathematics – it should be emphasized that the decision asserts the identity of those two discourses, not between mathematical objectivity and being simpliciter – is a decision for which no justification can be offered except retroactively through the demonstration of its productive consequences for thought, the decision may be seen as invoking what might be described in Heideggerian terms as an epochal disclosure of being. That is to say, the Badiouian decision is compelled by the recognition that modern mathematics has enabled the categories of being, universality, truth, and subject to be rethought, and it is the task of philosophy to register the possibility opened by mathematics – a task that has a certain historical urgency for Badiou, who sees that the Platonist categories of truth, universality, and absolute have been denigrated to the detriment of thinking, from both within and without philosophy (Badiou 1999).

Showing that mathematics is the discourse of being does not require an extensive treatment of the most interesting or novel idea in mathematics, but only a demonstration that well-established practices in mathematics can be read as answering to the very same questions with which ontology as practiced by philosophers have struggled. The mathematical components of *Being and Event* almost exclusively consist of the standard Zermelo-Fraenkel set theory with the axiom of choice (ZFC) and a limited arsenal of mathematical techniques, with Badiou readily acknowledging that set theory is neither the most 'interesting' nor "significant in the current state of mathematics." (2005, 14) But if the proposition that mathematics is ontology is "not a thesis about the world but about discourse," (Badiou 2005, 8) taking Badiou's set-theoretical theorization of being as describing the *way things are* would elide the distinction that Badiou imposes between thesis about the world and discourse of being qua being. An ambiguity thus arises concerning what to do with the concepts elaborated in *Being and Event* once they are taken outside the system of rational immanence that Badiou's theoretical discourse is, and then applied to extraontological situations, such as

social or historical situations wherein any actual political processes must take place.

For any transposition of the insights of set-theoretical ontology that is elaborated in *Being and Event* to extraontological situations to be legitimate, it needs to be established, minimally, that those extraontological situations are such that their being could be thought under the intricate architecture of being qua being conceived within *Being and Event*. It is precisely this linkage, which pertains to what could be considered following Kant the *quid juris* question, proves difficult to conceive. Peter Osborne notes that Badiou's decision to sever ontology from "all phenomenological relations to objects" leaves him with "the awkward task of restoring the connection between his set-theoretical mathematical entities, philosophically received ontological concepts (like nature and history) and the world." (2007, 24) Peter Hallward similarly remarks that it is not clear whether there is a place for 'ordinary ontic reality' in the set-theoretical universe of infinite multiplicities on which Badiou founds his thinking of being (2008, 118). The problematic relationship, or lack thereof, between ontological and extraontological situations in Badiou's speculative philosophy leads Ray Brassier, a prominent interpreter of the Badiouian text, to pronounce that Badiou's philosophy "simply stipulates an isomorphy between discourse and reality, logical consequences and material causes, thinking and being. Thinking is sufficient to change the world: such is the ultimate import of Badiou's idealism." (2007, 113)

Ultimately, the question of whether Badiou's set-theoretical general ontology can be legitimately transposed to extraontological situations admits of no straightforward answer.[2] For this reason, studies that attempt to transpose Badiouian concepts onto domains beyond ontology have been forced to make an interpretive choice on the weight given to mathematics. In his *Badiou and Politics*, Bruno Bosteels expresses what is perhaps the most obvious choice:

> Metamathematical concepts are rigorously formal, and they hold true according to an intrinsic rationality only within the ontological situation; anywhere else, they are just helpful tools that by analogy, through a symptomatic reading of mathematical names […] or in a metaphorical transposition, may help us formalize situations that are not in and of themselves ontological in the strict sense. (2011, 35)

Bosteels opts to regard set-theoretical ontology as a sort of a helpful tool or a heuristic for thinking transformative processes that must unfold in situations that are not strictly ontological, asserting that "as soon as we exit the domain of strict ontology […] the role of mathematics becomes heuristic at best." (2011, 35) However, Badiou's own stance hints, without repudiating the passage from Bosteels, at a more intimate relation between mathematics (set theory, in particular) and extraontological situations:

[2] For an in-depth examination of the relationship between ontology and extraontological situations, see Tho (2008).

> We have a concrete situation. We can think the ontological structure of that situation. […] It is very difficult sometimes, but we can. So we can think about infinite multiplicity, something about the natural multiplicity, something about the historical character of the situation, something about the evental site and so on. There is an ontological schema of the situation. With this schema we can understand the situation. (Badiou 2004b, 178-79)

The reconstruction offered in this paper shall take its cue from the preceding statement from Badiou, namely, that to *think* the ontological structure of a concrete situation is to assume a certain perspective from which the situation is analyzed. On what that extraontological situation concretely is, set-theoretical ontology is utterly indifferent. For, just as semiotics as the viewpoint from which the totality of social life could be redescribed was made possible by the liberation of abstract system of formal rules governing the combination and substitution between elements from any particular substance that was still present in Saussure's concept of sign as the distinction between conceptual mass and acoustic substance (Laclau 2007), the formalism of set theory, its non-reference to a particular object, is what allows it to be the basis of a general ontology such that, while it may not grasp all the specific aspects of every region of being, no region of being – including that of the social – essentially resists its grasp.

3. The Structuration of Being in Badiou's Set-Theoretical Ontology

That *the one is not* is a fundamental metaontological decision of *Being and Event* (Badiou 2005, 23). Badiou stipulates that whatever oneness – determinacy or objectivity – there is, it is a result of what he calls an 'operation' of 'count-as-one.' The primary significance within Badiou's system of the initial positing of oneness as a result of some mechanism of unification is that it implies the necessity of positing not-one as anterior to the one-effect of the count-as-one. That which is posited as preceding oneness is called "inconsistency," described by Badiou as the "ungraspable horizon" of sheer being, which is qualified as multiplicity in as much as it is *not one* (2005, 34). While several prominent readers of Badiou have rightly noted a deep ambiguity in the notion of operation, whose operator, in fact, remains utterly anonymous within his strictly ontological considerations (something that shall be discussed in more detail later), its basic function is clear. The operation of count-as-one is an ordering into a minimal order of sheer being, and like the Levinasian category of the *il y a* to which it is thematically close, nothing can be said about it except that it must precede any determinate objectivity, identity, difference, and order. The resultant of a count-as-one is termed "consistent multiplicity" (which the ontological discourse inscribes with the strokes of { and }), or *presentation*. A *situation*, then, is defined as "any presented multiplicity," and Badiou proclaims there is "nothing apart from situations." (2005, 25) Thus, no intrinsic distinction between a presentation and a situation is stipulated by the ontological discourse. Nevertheless, a semantic distinction between the two can be proposed, as the distinction is useful especially

when the ontological discourse is transposed onto discourses on extraontological situations. Badiou remarks:

> Granted the effectiveness of the presentation, a situation is the place of taking-place [*le lieu de l'avoir-lieu*], whatever the terms of the multiplicity in question. Every situation admits its own particular operator of the count-as-one. This is the most general definition of a *structure*; it is what prescribes, for a presented multiple, the regime of its count-as-one. […] When anything is counted as one in a situation, all this means is that it belongs to the situation in the mode particular to the effects of the situation's structure. (2005, 24)

"To exist," for Badiou, is "to be an element *of*. There is no other possible predicate of existence as such." (Badiou and Hallward 1998, 130) And being an 'element,' he explains, "is not a status of being, an intrinsic quality, but the simple relation to-be-element-of, through which a multiplicity can be presented by another multiplicity." (Badiou 2005, 45) If a situation is a 'place of taking-place,' a kind of field of objectivity, then it shall be said that it is so in the following sense: a situation is an order – that is, a multiple that has its own regime (or rules) for presentations therein. The particular regime of count-as-one is the structure of a situation that makes a situation extensionally different from other situations. To exist is to be presented in a situation, which is also to say that whatever that exists in a situation *belongs* therein as an element. Allow it to be said, then, that every presentation, hence every objectivity, is situated.

Inconsistent multiplicity is not presented in a situation and therefore does not exist for the situation. But it cannot be the case that inconsistent multiplicity is simply non-being, given the philosophical stipulation that it is from inconsistency that any situation as a consistent multiple must have arisen. To phrase this in a Heideggerian manner: the occlusion of *Sein* –pure inconsistent multiplicity – in its disclosure as *Seiendes* – consistent multiples, presentations, situations – does not annihilate the former. In the structured presentation that a situation is, the inconsistent being from which it has emerged lingers as what Badiou describes as a "phantom remainder." (2005, 53) Inconsistency subsists in the situation as unpresentable. Every situation, in so far as it admits into itself the inconsistency that it 'unpresents,' is simultaneously, as Roland Végső (2013) had put it, a "failed presentation." This failure is not a failure to present something that could have otherwise been presented. Rather, it is a necessary and constitutive failure: in so far as a situation is always an operational result of count-as-one, no situation can be a situation without the subsistence of the inconsistency of which it is an operational result within itself. Because it is unpresented, inconsistency is *nothing* from the perspective of the situation – the inhabitants, so to speak, of the situation do not register inconsistency. But its unpresentable subsistence in a structured presentation can be thought, and be inscribed in the discourse of ontology, not as inconsistency as such (as it precedes any *set-ness*), but as the letter ∅, the empty set. To this letter, Badiou confers the name 'void.' The void thus designates the point through which a situation comes to be a particular situation

– hence the void marks the *suture* of a situation to its own being – but that which must be unpresented, thus count as nothing, for the consistency of the situation.

A unique characteristic of the empty set allows the key implications of the void to be thought under the ontological discourse. The empty set is a set that is included in (that is, it is a subset of) any set, even if it is not an element of (that is, presented in) that set. This characteristic allows incorporation in set-theoretical ontology the thought that although inconsistency is unpresented and excluded (or *subtracted*, in Badiou's terminology) from presentation (or from a situation), it is nonetheless simultaneously *included* as the underlying pure being of every presentation (or a situation). Although it is never locatable in a situation because it is unpresented, the void of that situation is nevertheless dispersed everywhere in that situation. In short: "insistence of the void in-consists as de-localization." (Badiou 2005, 77) In extraontological situations, what the void is for a situation will depend on what the situation is, or *of what* it is a structured presentation. Apropos social situations, Rancière's reflection on equality may serve to illustrate the point. According to Rancière, what is concealed by social distribution of places and functions is the simple equality of humanity qua beings with the capacity to speak rationally, on which he argues rests every social order and makes whatever hierarchical relations imposed within an order operative. Equality in this sense can be seen as the void of the social situation. It is universal in as much as it is empty, it is what remains when all the particular differentiating features of different individuals and groups are bracketed. Badiou puts it thus: "The law of the void is in-difference." (2005, 77) The void of any social situation can thus be termed 'generic humanity,' an infinite multiple that is not marked by any qualification other than being human. The void, in either case, is that from which any social order is woven, but simultaneously that which is subversive, thus must remain foreclosed from presentation if order – or the consistency of a social situation – is to be preserved.[3]

It had been mentioned earlier that every extraontological situation is doubly structured. One way to understand why Badiou proposes that a situation is structured again through a second count, a 'count-of-the-count,' is to regard the second count as an operation that makes the aforementioned failure of presentation to continue to remain undisclosed, such that the situation may persist in its oneness, as a consistent multiple. The second count reduplicates the oneness of a situation by re-counting, firstly, the presented elements of the situation, thus affirming the initial count-as-one by which those, and only those, elements are situationally presented. Secondly, the second count counts all the ways in which presented elements of the situation could be arranged, thereby rendering the situation complete or whole, assuring that nothing more can be

[3] It should be noted that Rancière himself rejects any 'transcendentalization' or 'ontologization' of equality. His proposition of equality, Rancière insists, is but an "opinion." (2009) Whether Rancière can maintain his ontological non-commitment consistently, of course, is an altogether different question.

made out of the situation. The count-as-one is the operation of situation. But what maintains the consistency of the situation by keeping inconsistency unpresented, hence what accounts for the sustenance of a certain regime of count-as-one, is the second count, the operation of what is called the "state of a situation." (Badiou 2005, 97) What comes to fore with the state (*état*) is the structuring of a situation beyond the minimum that it, conceived simply as structured presentation, itself does not impose. For example, a plate on which there is a variety of fruits is a situation whose regime of count-as-one is such that its presented elements are fruits. But seeing that there is nothing apart from certain number of fruits on the plate and categorizing those fruits into different types of fruits call for additional work – it is this additional work that is performed by the state. What the state of the situation, itself conceived as a set, admits into itself is not the presented elements of the situation as such – which pertains to *of what* the situation is an order – but subsets (or parts) of the situation. For a situation S, its state is its power-set $\wp(S)$. If S were a three-element set, {α, β, γ}, its subsets, in addition to ∅ (which is a subset of every set, thus universally included) and itself, would be the following: {α}, {β}, {γ}, {α, β}, {α, γ}, {β, γ}. Therefore, the set of all existing parts of S, $\wp(S)$, is: {∅, {α}, {β}, {γ}, {α, β}, {α, γ}, {β, γ}, {α, β, γ}}. About the element γ of the situation S, the following can be said: γ *belong* to S – this is to say that γ is counted-as-one in S; but γ is also *included* in S, in so far as it is re-counted by the state of the situation, which means: {γ} is an element of, hence belongs to, $\wp(S)$. A multiple (an element) that both belongs to and is included in a situation is said to be both presented and represented in that situation. In Badiou's typology, such a multiple is said to be 'normal.' Given Badiou's set-theoretical premises, it must also be said that any multiple that belongs to the state of the situation is a representation.

Already from the example of a three-element set, it can be observed that there are always more parts of a set than elements of that set. Set theory teaches that the cardinality of $\wp(S)$ exceeds the cardinality of the initial set S, and immeasurably exceeds, in the case that the initial set is an infinite set. And as Badiou stipulates that every situation is indeed 'ontologically infinite,' (2006, 143) the cardinality of the state of any situation immeasurably exceeds the situation of which it is the state. The excess of inclusion (representation) over belonging (presentation), however, leaves the door open for anarchy – if the number of possible arrangements of a situation is immeasurable, it defies the goal of the state, which is to render a situation complete. If the necessity of the state of the situation derives from the need to secure the consistency of the situation, it will need to police – the Rancièrian allusion is suitable here – the excess of representation by imposing certain constraints on the ways in which a situation could be ordered. It could be said, then, that the space of representation needs a *regime* of representation that ensures both that the void remain unpresented and the excess of representation over presentation be tamed. It may thus be concluded that every extraontological situation, in as much as it is an order and not chaos, is under the

influence of a regime of representation unique to it that maintains the situation as it is.

4. A Depiction of a Normalized Social Situation

While this paper must suppose that the ontological categories outlined in the preceding section can be productively mobilized to illustrate and understand social situations, it is the case that there is no intrinsic definition of 'social' that is derivable from the Badiouian general ontology founded on post-Cantorian extensional set theory. To proceed, then, a concrete situation that can reasonably be qualified as social must simply be selected. For the purpose of exposition, let it be supposed that 'Indonesia' names a social situation. This situation requires that there be at least one unique membership criterion which would ensure that it be a structured presentation extensionally different from other situations. There is at least one uniquely ascribable regime of count-as-one to the situation chosen, such that it will be possible to say that those that belongs to the situation, in as much as they belong, are Indonesian citizens. An inhabitant of the Indonesian situation would be placed in a situation where there are infinite number of multiples populated by other inhabitants of Indonesia (such as families, universities, baseball teams), which are themselves a multiple of multiples (a university, for example, will have classrooms, offices, and so on). The relationship between the social situation and the various presented elements encountered therein is that the latter are subsets, or parts, of the former. A baseball team, for example, would be a part of, or be included in, Indonesia if all the individuals that belong to the baseball team also happen to be Indonesian citizens, hence belong to the Indonesian situation as well. In this case, the baseball team is a presented multiple that is also represented, and the same can be said of all its individual members. In the Badiouian typology, the baseball team would be a 'normal' multiple.

The arrangement of the parts of a situation is regulated by the state of the situation, in accordance with a particular regime of representation. In his mathematically oriented reading, Burhanuddin Baki suggests that the state should be seen as "Badiou's version of the Lacanian symbolic." (2015, 105) The comparison is not unjustified: the particular arrangements of parts proceed through the deployment of an apparatus that Badiou calls the 'language of the situation' – to be explained in detail shortly – that mediates presentation and representation. Moreover, the goal of the state is to govern the relationship between presentation and representation in a particular way to produce an illusion of completeness or wholeness of the situation, so as to foreclose the situation's "encounter with its own void, the presentational occurrence of inconsistency." (Badiou 2005, 93) In this respect, the state of the situation is comparable also to the Lacanian imaginary, with the void, unnameable by the situation's language and foreclosed from presentation, analogous to the Lacanian real. The state of the situation regulated by a regime of representation is the

closest that the ontological discourse is able to offer as a schema for what could be termed a 'social imaginary,' conceivable as the set of representations of what the inhabitants of the situation can become, but whose actualization would not disrupt the situation's self-identical unity in so far as they can be anticipated from how the situation already is. The conceptual determination of novelty and change – the prime achievement of *Being and Event* – cannot be completed without thinking their very opposite, namely, the way in which a situation is maintained in its self-identical unity. Critical for understanding the situation's self-identical unity, in fact, is the language of the situation.

Along the lines of interpretation offered in this paper, language of the situation can be understood as the mechanism by which a regime of representation works to constrain the space of representation through two types of operation: discernment and classification. Discernment, Badiou explains, "concerns connection between language and presented or presentable realities," whereas classification concerns "the connection between the language and the parts of a situation, the multiples of multiples." (2005, 328) As Badiou's ontological discourse is extensional set theory, a property cannot be defined without prior existence of a multiple whose extension just is the extension of that property. Badiou regards as the materialist postulate of his ontology that "being is anterior to language" (2005, 501) – it is a prescription of set-theoretical ontology, ensconced in the axiom of separation, that "language cannot induce existence" and that "a predicate only determines a multiple under the supposition that there is already a presented multiple." (Badiou 2005, 44-47) This means, in turn, that the existence of at least one baseball team in the social situation guarantees an extension for the property of being a baseball player: it is simply all the individuals that belong to baseball teams. In so far as a property can be defined extensionally, that is, in so far as a property discerns a multiple in the situation, it is legitimate to mobilize that property, along with other definable properties, to then predicatively define some other set. This allows all presented multiples and inhabitants of the situation to be defined in terms of certain set of properties, hence classified. A regime of representation does not induce presentations, but names, defines, and classifies already presented multiples, and these determinations serve as the basis on which statements about the situation that can be verified as veridical or erroneous are made. Classifications might be based on gender, race, occupation, religion, and so on, but as every predicate in the language of the situation is liaised with presentation, every classification can be traced to whatever is already presented in the social situation, which, in the broadest sense, would be the material practices of human individuals. Multiples that are defined via the language of the situation, through operations of discernment and classification, are *constructible*. These multiples constructed through the workings of the situation's language provide the ontological schema of subject-positions or objective social identities.

The state of the situation of the social situation, S, is its power-set ℘(S), which is an infinite multiple whose cardinality is greater than the cardinality of S. But, importantly, in so far as it has as its elements subsets of S that are definable using the language of the situation, the state of the situation will only contain multiples that can be discerned and named in the situation with its language, its discursive resource, namely, multiples that are constructible. Thus, although the number of ways of arranging the situation – the space of representation –always exceeds, immeasurably, what is presented, it is possible for a specific regime of representation to keep the excess to a minimum by governing what enters into the space of representation. By means of the language of the situation, the state ensures a 'proximity' between presentation and representation such that "the state does not exceed the situation by *too much*, or that it remains commensurable." (Badiou 2005, 288) The imaginary of a normalized social situation is limited in the following sense: it does not go beyond what is already being said and done in the situation, in so far as the only kind of subset that it admits into itself always relates back to what is already being said and done in the situation. And from the stability of the situation, it is to be inferred that its regime of representation is effective in providing the inhabitants of the situation an experience of their world as orderly and intelligible.

Having defined the fundamental limitation of a normalized situation within the terms of set-theoretical ontology, it is possible to anticipate how change must be conceptualized: if there is a subset that is absent in the state of the situation regulated by a regime of representation, it would be a non-constructible subset of S, hence no predicate or combination of predicates of the language of that situation would circumscribe it. If the state of the situation regulated by a regime of representation is the schema of social imaginary, then that subset may be ineffable or even vaguely menacing to the inhabitants of the situation, who can, in any case, neither describe nor know what it really is. Such a subset is an *indiscernible* part of the situation. This part can, under certain conditions, be qualified as 'generic,' a term which Badiou also characterizes the being of what he rather controversially calls a 'truth' – it is the key objective of the second half of *Being and Event* to advance the idea that radical change in the situation is induced by its supplementation by a truth. Something beyond imagination can only appear as an unpredictable, aleatory occurrence from the perspective of the situation. Hence the centrality for Badiou's theorization of change the concept of event, to which the second half of the title of his treatise refers. But the final section of this paper must leave aside the theory of change proper and return to consider an important feature of pre-evental social situations whose fundamental structures have been described using the resources of Badiou's set-theoretical ontology.

5. Social Situations without Subjects

As hinted earlier, the ambiguous status of the anonymous 'operation' by which presentations and representations result, the count-as-one, has been noted by

several commentators. In one of the earliest substantive reviews of *Being and Event*, Jean-Toussaint Desanti turns to Badiou's following declaration: "What has to be declared is that the one, which is not, solely exists as *operation*. In other words: there is no one, only the count-as-one." (Badiou 2005, 24) Desanti proceeds to write: "the project of a pure ontology (an intrinsic theory of being as being) would stumble here with its very first step, were one to ask oneself this 'preliminary' question: what is it to operate? Who operates here and in what realm?" (2004, 60) Edoardo Acotto, similarly, asks: "Badiou defines [the count-as-one] as an *operation*. But who, concretely, is the *operator*? This is one of the mysteries of Badiou's philosophy, and of its exclusion of perceptual and cognitive mechanisms from the ontological discourse." (2007, 86; translated from French)

The question of the 'counter' in the operation of count-as-one (to which a sort of Kantian transcendental subjectivity cannot be an acceptable answer for Badiou) challenge Badiou's philosophy – apart from demanding a more robust account of the transition from inconsistency to consistency – to account for the genetic question of how *particular* situations come to be. A question that is also relevant, especially for thinking social situations, is that of *to what* particular social situations that result from regimes of counting-as-one owe their *force* in persisting as they are. As Paul Livingston has argued, irrespective of what Badiou intends to achieve philosophically (such as avoiding recourse to subjective or linguistic idealism), at least apropos social situations, it is "very difficult to avoid the natural assumption" that what is operative in the persistence of their presentational and representational regimes ultimately alludes to "structures of linguistic or conventional practice, established and held in place by the behavioural regularities of a specific cultural or language community." (2014, 245) To be sure, it may be that for set-theoretical ontology to be a general ontology, its account of structuration cannot be inextricably tied to any particular counter, or that a general ontology founded on set-theoretical axioms is able to consistently deny any intrinsic definition of what a situation is, thus refrain from providing a general condition for *what counts as a situation* and why. Yet, as soon as the specificity of particular situations or types of situations is considered, it seems to become critically pertinent to ask "[w]hat is to count as a situation, and who decides." (Eagleton 2003, 252) Apropos social situations, Badiou himself appears to be willing to allow linguistic and conventional social practices to be implicated in their presentational and representational structures, as attested by several of Badiou's own examples. If, however, it is the case that operations of counting, at least in social situations, could be understood in terms of linguistic or conventional practices, then the conspicuous absence in the account of social situations of any sort of agent or some form of subjectivity that such practices likely need to presuppose appears as a problematic omission.

The depiction of social situations as asubjective, which remains as the only depiction that is directly supported by the Badiouian ontological discourse, can be contrasted with accounts offered by theorists that give greater prominence to the

role of discursivity in the constitution of the social, such as Ernesto Laclau, whose final publication is tellingly titled *The Rhetorical Foundations of Society* (2014). In the entirely asubjective action of structure postulated by the discourse of set-theoretical ontology and an understanding of social situations based on it, there seems to be little room for the thought, important in Laclau's approach, that there is a distinction between identification and identity that corresponds to the distinction between the moment of the subject and that of subject-position, instituting an order and an instituted order, or between the political and the social. In the set-theoretical account of social situations, it seems that human inhabitants are passively captive to anonymous operations of counting, hence relegated to the status of structurally determined objects. For Laclau, however, a structure cannot fully constitute its elements, as well as the structure itself, as self-identical objects. He contends that it is precisely this failure of 'structural objectivity' that opens the 'space of the subject.' (Laclau 1990, 67) In other words, it is because social agents construct and identify with certain subject-positions as an attempt to overcome the failure of structure that a social order is configured and reconfigured in some particular way. From this perspective, subjectivity is always implicated in how a social situation is structured and restructured. In the set-theoretical account, however, neither such creative acts of identification nor whatever that compels such acts – for instance, the subject's desire that Lacan-influenced theorists propose to understand with reference to the idea of the ever-elusive object of desire, the *objet petit a* – has a place.[4] Yet, it would still be too hasty to conclude its exclusion of such subjective acts makes it impossible for Badiouian ontology to support a robust social ontology, without considering a rationale that can be offered from the Badiouian perspective for the absence of consideration of human agency in normalized, pre-evental social situations.

One of the key objectives of *Being and Event* can be said to be the conceptual determination of processes by which "situations *necessarily* transform themselves to accommodate the existence of something that had not been acknowledged until that point." (Gillespie 2008, 79) Badiou's system imposes a strict distinction between modifications that occur as an immanent development of the situation, which is regulated by the state, and a real change that begins from the interruption of that development. The former kind of change is one that Badiou associates with the constructivist orientation. It is indeed this orientation that has been adapted in this paper to depict normalized situation. In the constructivist orientation, "[w]hat is called 'change' in a situation is nothing more than the constructive deployment of its parts. […] A new nomination takes the role of a new multiple, but such novelty is relative, since the multiple validated in this manner is always constructible on the basis of those that have been recognized." (Badiou 2005, 290) The type of change that occurs in a normalized situation is – to repeat Cornelius

[4] On the use of Lacanian psychoanalysis in Laclau's political theory, see Stavrakakis (1999) and Stavrakakis (2007).

Castoriadis's critique of classical structuralism that captures precisely the gist of the constructivist orientation – "no more than a spreading out" of the situation, such that what counts as "the new is, in every instance, constructed through identitary operations by means of what was already there." (1997, 173) Within normalized situations, identifications harbor no transformative potential regardless of the 'agent' behind them, in so far as they merely unfold in accordance with the representational regime of respective situations. Attempts to produce 'new' identities or subject-positions are applications of the discursive resources of a situation – its language – to produce additional constructible multiples, where such productions do not necessarily compose a process that brings something new into the situation, in as much as they are redeployments of what is already presented in the situation and to what already is being said and done. This renders identification and productions of identities in themselves irrelevant for the theorization of change as the emergence of new, unanticipated possibilities in a situation, as such a theorization requires a strict conceptual distinction between that which is actually new and that which is only possibly new.

The spatial metaphor of 'spreading out' in the line from Castoriadis is appropriate to describe the kind of change possible under a normalized social situation, which is not exposed to something other than what it already is or what is anticipatable from within it – normalization in this respect is the spatialization of temporality, and a normalized situation is one in which the synchronic prevails over the diachronic. Constructivist in orientation, the regime of representation and its discursive apparatus cannot be the source of novelty or induce radical change. There is – to repeat Rancière's apt remark on the policed distribution of social roles – a sense in which that "everyone is included in advance" in a normalized situation (1998, 116). Whatever 'new' identity definable using the situation's language will never fail to be represented. Because the subject for Badiou is "the real presence of change in a situation, or the actual existence of the new," (Pluth 2010, 118) there is no subject prior to an event. Subjectivation, the emergence of a subject in the situation, is subsequent to the interruption of a normalized situation in an event, whose conceptualization, Badiou contends, is beyond the remit of ontology. For, in as much as an event is not of the order of being but signifies its radical interruption, it is – as Emmanuel Levinas, with whose ethical philosophy Badiou shares more in common than he would be willing to admit, might put it – *otherwise than being*.

6. Conclusion

If the pre-evental social situation reconstructed from Badiou's set-theoretical ontology appears to be a monotonous world in which nothing genuinely new takes place, this is in no small part due to the fact that, within the architecture of Badiou's system, the depiction of the pre-evental situation is the propaedeutic to a dichotomic determination of what change and socially transformative political processes are. From the perspective of this system, events and processes of real

change are far rarer than would be supposed by a worldview that legitimates the extension of the referent of 'activism' from the factories to the streets to social media hashtags. While the ascription of rarity to change could be seen as indicative of a problematic yearning for a 'great politics,' it may simultaneously be a sobering reminder that – as Badiou had remarked apropos the *gilets jeunes* movement in France – "not everything that moves ahead is red [*tout ce qui bouge n'est pas rouge*]." (2019)

Despite its purported aim of theorizing novelty and change, whether Badiou's intricate system actually is able to support the thought of a far-reaching social change, particularly in the direction endorsed by Badiou himself, is an open question. Difficulties that arise when conceptualizing social change within the coordinates of Badiou's system have been documented by a thought number of readers. Livingston's elaboration of those difficulties stands out for its attentiveness to the implications of set-theoretical ontology (2014). Laclau, for his part, has challenged the strict dichotomy Badiou imposes between the pre-evental situation and post-evental processes of genuine change (2004). Commentators have also argued that there is no reason to believe that any particular event and post-evental change should command anyone's approval and commitment – a normative deficit, as it were, haunts Badiou's "ethics of truths." (Dews 2004; Lecercle 1999; Vilaça 2014) Perhaps, even Badiou's appropriation of the word 'truth' to name a crucial category in his philosophy may be challenged, in as much as Badiou fails to tie truth sufficiently to a notion of the good (Critchley 2000; 2007).

Difficulties such as those aforementioned have not, however, prevented the application of Badiou's account of situations and change to social theory. The most extended of such attempts to bring Badiou's thought to bear on social theory to date is Colin Wright's *Badiou in Jamaica* (2013). Wright demonstrates in his study that the Badiouian system presented in *Being and Event* and *Logics of Worlds* can be further elaborated into a sophisticated conflict theory, while also entering into a productive dialogue with experientially and historically grounded thinkers from cultural studies and postcolonial studies. The potential of Badiou's system to contribute to social theory cannot be measured in advance. Applying the same criterion by which, according to Badiou, events are to be ultimately evaluated, it might be said that a great portion of the evaluation of Badiou's thought will depend on the consequences it produces, both in theory and in practice. The reconstruction of the social ontology contained within *Being and Event* in this paper has been offered with the hope of continuing their production.

References

Acotto, Edoardo. 2007. "Alain Badiou et l'ontologie du monde perdu." In *Écrits autour de la pensée d'Alain Badiou*, edited by Bruno Besana and Oliver Feltham, 83-100. Paris: Editions L'Harmattan.

Badiou, Alain. 1999. *Manifesto for Philosophy*. Translated by Norman Madarasz. Albany: State University of New York Press.

————. 2004a. "Being and Appearance." In *Theoretical Writings*, translated by Alberto Toscano and Ray Brassier, 163-76. London: Continuum.

————. 2004b. "Ontology and Politics: An Interview with Alain Badiou." In *Infinite Thought: Truth and the Return of Philosophy*, edited by Oliver Feltham and Justin Clemens, 169-93. London: Continuum.

————. 2005. *Being and Event*. Translated by Oliver Feltham. London: Continuum.

————. 2006. *Metapolitics*. Translated by Jason Barker. London: Verso.

————. 2009. *Logics of Worlds*. Translated by Alberto Toscano. London: Continuum.

————. 2018. *L'Immanence des vérités*. Paris: Fayard.

————. 2019. "Lessons of the 'Yellow Vests' Movement." Translated by Gregory Elliott. Verso. Accessed March 20, 2021. https://www.versobooks.com/blogs/4327-alain-badiou-lessons-of-the-yellow-vests-movement.

Badiou, Alain, and Simon Critchley. 2007. "Ours Is Not a Terrible Situation." *Philosophy Today* 51 (3): 357-65.

Badiou, Alain, and Peter Hallward. 1998. "Politics and Philosophy: An Interview with Alain Badiou." *Angelaki* 3 (3): 113-33.

Baki, Burhanuddin. 2015. *Badiou's* Being and Event *and the Mathematics of Set Theory*. London: Bloomsbury.

Bosteels, Bruno. 2011. *Badiou and Politics*. Durham: Duke University Press.

Brassier, Ray. 2007. *Nihil Unbound: Enlightenment and Extinction*. New York: Palgrave Macmillan.

Castoriadis, Cornelius. 1997. *The Imaginary Institution of Society*. Translated by Kathleen Blamey. Cambridge: The MIT Press.

Critchley, Simon. 2000. "Demanding Approval: On the Ethics of Alain Badiou." *Radical Philosophy* 100: 16-27.

————. 2007. *Infinitely Demanding: Ethics of Commitment, Politics of Resistance*. London: Verso.

Desanti, Jean-Toussaint. 2004. "Some Remarks on the Intrinsic Ontology of Alain Badiou." In *Think Again: Alain Badiou and the Future of Philosophy*, edited by Peter Hallward, 59-66. London: Continuum.

Dews, Peter. 2004. "States of Grace: The Excess of the Demand in Badiou's Ethics of Truths." In *Think Again: Alain Badiou and the Future of Philosophy*, edited by Peter Hallward, 106-19. London: Continuum.

Eagleton, Terry. 2003. *Figures of Dissent: Critical Essays on Fish, Spivak, Žižek and Others*. London: Verso.

Gillespie, Sam. 2008. *The Mathematics of Novelty: Badiou's Minimalist Metaphysics*. Melbourne: re.press.

Hallward, Peter. 2008. "Order and Event." *New Left Review* 53: 97-122.

Laclau, Ernesto. 1990. *New Reflections on the Revolution of Our Time*. London: Verso.

————. 2004. "An Ethics of Militant Engagement." In *Think Again: Alain Badiou and the Future of Philosophy*, edited by Peter Hallward, 120-37. London: Continuum.

————. 2007. "Discourse." In *A Companion to Contemporary Political Philosophy*, edited by Robert E. Goodin, Philip Pettit, and Thomas Pogge, 2nd ed., 1: 541-47. Oxford: Blackwell.

————. 2014. *The Rhetorical Foundations of Society*. London: Verso.

Lecercle, Jean-Jacques. 1999. "Cantor, Lacan, Mao, Beckett, *même combat*: The philosophy of Alain Badiou." *Radical Philosophy* 93: 6-13.

Livingston, Paul M. 2014. *The Politics of Logic: Badiou, Wittgenstein, and the Consequences of Formalism*. New York: Routledge.

Maniglier, Patrice. 2010. "The Structuralist Legacy." In *After Poststructuralism: Transitions and Transformations*, edited by Rosi Braidotti, 7: 55-81. The History of Continental Philosophy. Chicago: University of Chicago Press.

Osborne, Peter. 2007. "Neo-Classic: Alain Badiou's *Being and Event*." *Radical Philosophy* 142 (April): 19-29.

Pluth, Ed. 2010. *Badiou: A Philosophy of the New*. Cambridge: Polity.

Rancière, Jacques. 1998. *Disagreement: Politics and Philosophy*. Translated by Julie Rose. Minneapolis: University of Minnesota Press.

————. 2009. "A Few Remarks on the Method of Jacques Rancière." *Parallax* 15 (3): 114-23.

Renault, Emmanuel. 2016. "Critical Theory and Processual Social Ontology." *Journal of Social Ontology* 2 (1): 17-32. https://doi.org/10.1515/jso-2015-0013.

Stavrakakis, Yannis. 1999. *Lacan and the Political*. London: Routledge.

————. 2007. *The Lacanian Left: Psychoanalysis, Theory, Politics*. Albany: State University of New York Press.

Tho, Tzuchien. 2008. "The Consistency of Inconsistency: Alain Badiou and the Limits of Mathematical Ontology." *Symposium* 12 (2): 70-92.

Végső, Roland. 2013. *The Naked Communist: Cold War Modernism and the Politics of Popular Culture*. New York: Fordham University Press.

Vilaça, Guilherme Vasconcelos. 2014. "Badiou's Ethics: A Return to Ideal Theory." *Badiou Studies* 3 (1): 271-321.

Wright, Colin. 2013. *Badiou in Jamaica: The Politics of Conflict*. Melbourne: re.press.

"A Superior Anthropological Perspective." On Kant's Anthropo-cosmological Conception of Ideal

Fernando Silva

Abstract: The topic of the ideal, that is, the topic of the *possible or impossible human attainment of the absolute* is ascribed divergent treatments throughout Kant's work. Namely, it is either promptly accepted as possible by the critical Kant, and seen as something attainable by *a means other than* an infinite approximation (which would indeed imply a violation of autonomy, but denies the genuineness of the ideal), or it is rejected as impossible by the non-critical Kant, that is, it is seen as something attainable *only through* an infinite approximation (which would involve an unconditional acceptance of heteronomy, but safeguards the authenticity of an aspiration to the ideal). Yet, the topic of the ideal receives a new, if not conciliatory, at least mutually explanatory approach in Kant's Anthropology. Here – such is our proposition – Kant proposes a *terminus medius* between both conceptions of ideal, insofar as he is led to ponder on the mutual benefits of an autonomic possibility and an heteronomic impossibility of an infinite progression in thought; something which Kant proposes under the form of an almost-infinite, or an almost perennial, yet finite duration, to be endured until the attainment of an almost unreachable, yet indeed reachable *practical ideal*. A *terminus medius* which, we hope to prove, is none other than that at the root of Kant's proposition of Pragmatic Anthropology as a mediating science in Kant's fundamental scheme of human knowledges, and which therefore may be ultimately seen as the embodiment of Kant's anthropo-cosmological, or indeed cosmopolitical dimension of thought, as expressed in Kant's political and/or historical writings.

Keywords: Anthropology, cosmopolitanism, Immanuel Kant, Metaphysics, practical ideal.

I. Introduction. The Ideal as a Kantian Problem

A core issue since the beginning of philosophical thought, the problem of the ideal, or the infinite, or the absolute, and its possible or impossible attainment by the human spirit, has been the subject to the most distinctive approaches by the most heterogenous authors. Now arising as Plato's union between finite and infinite,[1] as the unfathomable eternity, as in Haller (1882, 149-154), or even as the unreachable concept of totality – of the *hen kai pan* – among German idealists, the problem has given rise to important and no less fertile debates, and has also helped shape the history of philosophy – all the more, since from this issue

[1] As visible in *Timaeus* and analyzed by F. W. J. Schelling in one of his youth texts, "Über den Geist der Platonischen Philosophie." (Franz 1996, 282-320, Appendix II)

partially depend the boundaries, the task and destination of philosophical thought itself. One of the periods when this issue became most crucial, serving as the reflective propelling wheel of a whole philosophical era, was between the so-called Leibniz-Wolffian school and the German Idealism and Romanticism: a century during which the topic, until the first half of the 18th century handled as a strictly philosophical and/or metaphysical problem, outgrew philosophy and became a focus of discussion in Aesthetics, Literature, Psychology and Anthropology. As such, on the one hand, Leibniz, Wolff, all dealt with infinitude as an exclusively metaphysical problem, and in unison accept Man's proneness to tackle by thought – to perceive – the ideal of human knowledge, as well as to attain rational ideals; on the other, not one century later, young idealists such as Erhard, Novalis, Hölderlin or Forberg, conversely negated this metaphysical vision of the problem, rather asserting Man's possibility to attain the absolute, that is, the possibility of an infinite or endless approximation to an ideal;[2] and, in between both parties, stand those authors who, each in their own way, promoted a slow yet sure transition and fertile commerce from one side of the question to the other: singular authors such as Hamann, Baumgarten and, last but not the least, *Kant*.

Kant's treatment of the problem of the ideal – the focus of this article – is, in our view, evocative of such a transitive and expansive period in this private history of a philosophical problem, and this in more than one way. He himself an advocate of Man's necessary attainment of rational – regulative – ideals in his metaphysical writings, as well as, consequently, an accuser of Man's impossible attainment of aesthetic, historical, psychological – constitutive – ideals in his non-philosophical writings, Kant, however, did not limit his positions on the ideal, as above, to his Metaphysics, rather reexamined the problem from several other prisms, and reapplied such prisms, and their beneficial consequences, to several other dimensions of his work. Freeing himself from the limitation of having to solve the problem either by its unilateral negation – the primacy of the ideal over the human – or by its passionate acceptance – the opposite of the latter – it is our

[2] See Johann F. Hölderlin, "an infinite approximation (…), as is the approximation of the square to the circle." (1966-1969, 181) See also Novalis, "This absolute given to us is to be known only negatively, insofar as we act and find that what we seek is attained through no action." (1999, 181), Johann Benjamin Erhard: "This procedure of searching for the supreme grounds through reflections, of correctly subsuming the consequences and comparing the concordance of these consequences with that which was proved correct in experience, and of deeming this as true only in face of its harmony with experience (…), was to this day the method of the sane human understanding, and it is advisable that we abide by it until philosophical adepts are granted the opportunity to find the philosopher's stone or the supreme principle, of which all truths are to be weaved as from a ball of thread." (1970, 10) and Friedrich Karl Forberg: "That speculative reason is in want of a supreme principle, from which all truths are to be weaved as from a ball of thread, is undeniable. But I fear that what happened to alchemists, with their philosopher's stone, shall happen to philosophers and their supreme principle. They shall search for it endlessly and shall never find it. This is a task that nature gave reason not so that reason finds its resolution, but so that reason searches for it." (Frank 1998, 452-453)

view that upon laying the first bricks towards the foundation of Anthropology, Kant indeed inaugurates *a new hybrid, multi-relational field, a new discourse on Man and a new vision on Man and his conduct in the World*. But, not unrelated to this, and yet much beyond this, by paving the way towards a new comprehension of Man's position and task in the universe, Kant at the same time inaugurates a new alternative, much neglected[3] standpoint, as well as a singular perspective over the problem of Man's relation to the ideal, devoid of the unilaterality of the other two positions. Namely, imbued with an anthropo-cosmological, or practical vision of Man's earthly destination, Kant proposes a finite, yet extremely remote course of Man's formation, to which concur his ever growing historical, political, moral and cosmological self-cognition and action, thereby creating space for a new vision of human ideals in general.

As such, the present article envisages a double objective.

Firstly, to address Kant's account of the problem of the ideal – here under the guise of the finite and the infinite – in his early *Lectures on Metaphysics*. The aim is to ascertain the essential traits of the problem between the possibility and impossibility of experiencing the ideal, and how Kant's metaphysical view of this faces a necessary dilemma.

Secondly, to contrastingly analyze Kant's position on the problem of the ideal in the *Lectures on Anthropology*. The aim here is to show how, by dealing with the problem of Anthropology's position and task amid other akin and unfamiliar sciences, and by founding it in special, multi-relational connections with the latter, Kant simultaneously gains way towards a new cosmopolitical conception of an ideal of human knowledge.

II. The Problem of the Ideal from a Metaphysical Point of View

II. 1. Kant's *Conceptions of Infinite* and Their Link to a Fundamental Division of Human Knowledges

As is known, Kant initiated his *Lectures on Metaphysics* with a fundamental scheme of human knowledges: more specifically, a scheme depicting the division, and respective opposing disposition, of the different human knowledges.[4]

[3] Kant's concept of ideal, though not recurrently approached, has received some attention in all dimensions of Kant's thought. On this, see: Allison (2004) (especially the chapter "The Ideal of Pure Reason"); Schneider (2019); Zöller (1991, 52-59); Kleingeld (2012); Gerwen (2009, 331-345). However, the suggestion of a third, cosmopolitical dimension of knowledge, endowed with its own specific manner of progression – that is, its own conception of the (un)attainability of the ideal – as far as we know, has not yet been contemplated, not even mentioned, in any of the aforementioned works. As such, we present it here as our own new contribution to the field, in the hope that it will be able to complement the previous studies on Kant's conception of ideal.

[4] Such a scheme resurfaces *Metaphysik-Mrongovius* (Kant 1980, 747-940), *Metaphysik-Volckmann* (Kant 1968, 355-459), *Metaphysik L2* (Kant 1968, 525-610) or *Metaphysik-Dohna* (Kant 1968, 615-702), as well as in the *Lecture on Encyclopedism* (Kant 1980, 5-45) and the

According to this scheme, in Met.-Pölitz I, all knowledges are either an *aggregate*, or a *system*, and hence are opposed: "all knowledges are united either through coordination, or through subordination." (Kant 1968, 171) Namely, according to Kant, knowledges are *subordinated* because they proceed as if they were on a "ladder." (*Leiter*) (1968, 171) Subordinated cognitions compose a dimension of knowledge deemed "profound or grounded knowledge," (Kant 1968, 171) which cannot interfere with experience, and hence, this dimension is *finite* when considered regarding "the limits of human knowledge, which the human understanding cannot supplant." (Kant 1968, 171) On the other hand, knowledges are *coordinated* when they "proceed amongst themselves as parts in relation to a communitarian whole." (Kant 1968, 171) These relate as if they were on a "plain soil" (*ebenen Boden*) (Kant 1968, 171) and compose a different dimension of knowledge deemed "extended knowledge," (Kant 1968, 171) which diverges from rational knowledge and its principles by being *infinite*; for although "some sciences have limits that are determined by Nature and by reason itself" (Kant 1968, 172) – as do subordinating sciences – the limits of other sciences, however, "depend on the free will of men." (Kant 1968, 171) And so, in a word, it is Kant's view that given such a diametrically opposed divergence, the 'ladder' of rationality and the 'plain soil' of empirical knowledge form two lines of knowledge: one, a vertical, unitary, *finite*, line characteristic of intellectuality or rationality, which pervades the whole alignment of sciences from the least rational of the sciences of applied Metaphysics to the supreme transcendental philosophy; and another one, an horizontal, *infinite* line of experience, in time and space, engulfing all historical sciences, such as Empirical Physics, Empirical Psychology, or History.

Kant's conception of finite and infinite, which indeed is enrooted in this fundamental scheme, and therefore blossoms in the Prolegomena to the *Lectures on Metaphysics*, is however duly expounded only shortly after, in the same lectures, namely, in the sections on the finite and the infinite, in the part on Ontology.[5] This conception, (not without surprise) expounded in terms obedient to those of the previous fundamental scheme of human knowledges, presents the question of the appreciation of greatness, first and foremost, as a question of opposition between two different modes of consideration. The opposition, according to Kant in Met.-Volckmann, is one between quality and quantity ("per qualitatem qualitatis") (Kant 1968, 439) Namely, the concept of whole (*toti, Ganze*) presupposes a qualitative concept of totality (*omnitudo*), whereas the concept of quantity (*quanti, Menge*) presupposes a quantitative concept of maximum (*maximi*).

According to Kant, on the one hand, "the maximum is a relative concept," for it "gives me no determinate concept" (1968, 568) of what maximum that

much contested *Lectures on Geography*, *Geographie-Rink* and *Geographie-Vollmer*, unpublished in the *Akademie-Ausgabe*.

[5] Namely, "De progressu et regressu in infinitum," in Met.-Pölitz I (Kant 1968, 197-200); "Vom Endlichen und Unendlichen," in Met.-L2 (An Pölitz 3.2) (Kant 1968, 568-569); "Vom endlichen und unendlichen," in Met.-Volckmann (Kant 1968, 438-440).

maximum is. That is, quantity, because it is an extensive concept – "quantitas est vel aggregate, dicitur extensiva" (Kant 1968, 438) – proceeds by aggregation, or addition, which results in the eternal possibility of ever adducing something else to the line of addition (see Kant 1968, 568), and hence rendering the maximum ever more maximum. On the other hand, "the totality is an absolute concept," (Kant 1968, 568) for it provides the final unity of the envisaged whole. That is, totality, because it is an intensive concept,[6] proceeds via intensification, via 'degrees,' through 'distension,' and therefore greatly differs from the quantitative, which is an "expanded greatness" (Kant 1968, 438) – which results in a real and immediate possibility of attaining the totality. This contrasting view means that the manner of conception of the totality is qualitative, for totality itself is a *system* of degrees, or intensities, from the weakest to the strongest, but impervious to alteration; and the manner of conception of the maximum is quantitative, for the maximum itself is an *aggregate* of objects, of apparently neutral cognitive traits which, for this reason, are always exposed to further addition or subtraction. And so, Kant states, as many as the manners of conceiving greatness, there must be *two meanings of infinite*, which are distinguished insofar as they present themselves as the possible or impossible course of Man towards their obtainment: a first one, which is "a pure concept of understanding," (1968, 568) is deemed by Kant as an "infinitum reale," or "infinitum metaphysicum" (1968, 439): an infinite devoid of limitations or negations (see Kant 1968, 568), inasmuch as it contains the determinate concept of itself, and hence "contains all reality." (1968, 439) A second one, "referring to space and time" (1968, 568) – concludes Kant – is the "mathematical infinite," "which arises through the successive addition of one to one." (1968, 568-9) This, which indeed is subjected to limitations or negations – for lack of a determinate concept – "does not possess all reality" (Kant 1968, 439) and is eternally incomplete. And so different are the two infinites, that they can only correspond to two human modes of assessing greatness between which there can be no connection, much less interference: namely, the *qualitative vision of totality*, because it is a closed totality and of a rational nature, is *finite* (hence the infinitum is real), and the *quantitative vision of the maximum*, because it is an eternally open quantity and of a spatial and/or temporal nature, is *infinite* (hence the infinitum is mathematicum).

Now, thus collocated our problem, the consonances between this position and Kant's fundamental scheme of human knowledges are too evident to be a mere contingency. For, in view of the two structuring lines of human knowledge, and especially their opposition, as well as the different conceptions of their limits, or lack thereof, it cannot be a coincidence that, upon assessing the possible maximum, or total greatness of such knowledges, Kant states that the "infinitum reale" or "metaphysicum," a "mere concept of pure understanding" (1968, 438)[7] –

[6] "The greatness of that which is immediately represented as unity is intensive greatness." (Kant 1968, 438)

[7] See also Kant 1968, 568.

just as the rational 'ladder' in which it dwells – proceeds by subsumption, or by degrees of intensity and is an impermeable rational totality. Nor can it be a mere coincidence that Kant refers that, conversely, the 'infinitum mathematicum,' which refers to space and time – just as the 'plain soil' in which it dwells – proceeds by addition and is an ample, i.e., infinite aggregate. Indeed, we affirm, *these are no mere coincidences*. Quite on the contrary, *the 'infinitum reale' and the 'infinitum mathematicum' of human knowledge are, or constitute themselves, the above described rational and empirical lines of human knowledge and their sciences*; and, likewise, the vertical, or rational, and the horizontal, or empirical lines of knowledge encapsulate, or come to be formed in light of such infinites. This means, then, that Kant proposes a deliberate application of the fundamental methodological traits of his metaphysical reflection to one of its key-topics; just as he proposes a reapplication of this key-topic to the task and the destination of Metaphysics; a fact which, needless to say, cannot but have important implications to the development of the question.

As such, we say, already setting out in search for new implications: the problem at hand, more than just a problem of finitude or infinitude, and more than just a problem of the different natures of knowledge, is a problem of the possible limits, or possible non-constriction, of human knowledge in general. Namely, just as there are two kinds of knowledge, and two kinds of associated sciences – rational sciences and historical knowledges – so too are there two different procedures of the scientific approach in ascertaining the greatness of those knowledges – one could say, two scales of (human) measurement of knowledge: one reachable, the other unreachable; one bearing in mind the infinite of human reason, one bearing in mind the infinite of the World. Or, to put the question into more solid terms: one, the 'infinitum reale,' by which Kant alludes to the real – because finite – possibility of Man ascending in the ladder of rationality, thus attaining, through the metaphysical sciences in general – and, to be sure, transcendental philosophy – *the ideal(s) of reason*; the other one, the 'infinitum mathematicum,' in which Kant refers to the mathematical – for infinite – impossibility that Man progresses until the end of the ample field of historical knowledges, and hence his impossibility to attain the ideal, any ideal, by this means – *the ideal(s) of experience*. And if we think that, upon proposing the previous scheme of his metaphysical thought, Kant defends that such lines of human knowledges stand as dissociated, then, because one problem must reflect the other – *is the other* – exactly the same is to be applied to the (in)finitudes proposed: namely, they too, just as the knowledges which nourish them, are in irreconcilable separation from each other.

And so, we ask: how should we definitively enunciate the difference between such modes of measuring the maximum, or the totality, of human knowledges – now in their connection with the ideal? Apparently, by presenting these two as completely opposed: and this with different repercussions upon the problem of the ideal.

Namely, on the one hand, the 'infinito metaphysico,' which crosses the whole rational 'ladder' of human knowledge, is a "defined greatness" (Kant 1968, 439); that is, it is lesser than any number, and hence it is (at least) thinkable to the human understanding. In a word, one could say, the ideal is here a pure concept of the understanding in (or for) the understanding itself, and hence attainable. And so, one could say, the metaphysical infinite proposes a finite ideal, as finite are rational sciences; or, in other words, *an ideal which is not exactly an ideal*, for we do not have to aspire to it, not at least in the sense that its cognition may offer resistance, may hide from the understanding and be inaccessible to it (for, so Kant, it does not have negations, nor limitations). Of this sort are, for instance, the ideals of pure reason.[8]

Conversely, the mathematical infinite, or the ideal of the historical line of knowledge, is based upon the 'plain soil' of knowledge: a fruit of the negations and limitations that characterize it, and mark its infinitude, it offers the human understanding a resistance and is ideal by its own right. But, because the mathematical infinite is centered upon a "omnitudem collectivam," (Kant 1968, 569) because it attempts to apprehend a "totality of phenomena," (Kant 1968, 569) in a word, because it tends to assess and be a greatness that is"'greater than any number" ("greater than that which I could ever measure") (Kant 1968, 439)[9] then it offers us a never entirely surmountable, for never entirely terminal, resistance. This, one could add, represents an 'infinite progression,' or an *infinite approximation* to the ideal, which, according to Kant, "is never to be thought" (1968, 440) and is forbidden to the human understanding: "The totality in the collective infinitude, which is opposed to the progressive – which is already given as infinite – must have consummated an infinite progression, and this is not to be thought; for in this precisely consists eternity, that the progression can never be consummated (…)" (1968, 440).[10]

---

[8] See Kant 1911, 383-391.

[9] "In the first one, infinito metaphysico, we have a determined greatness, for it contains all reality, whereas in the mathematical infinite we have no determinate greatness, rather we know only that it is greater than any number. (…) Hence, from mathematical infinite space I know nothing apart from the fact that it is greater than that which I could ever measure, for we know everything only through addition, and if this is not possible, then this surpasses all our means to know such a thing, hence I cannot cognize it entirely in relation to my understanding." (Kant 1968, 439)

[10] See Kant's position on aesthetic ideals in the third *Critique*, namely, "the impossibility of the absolute totality of an unending progression." (1913, 255) The topic is once again resumed, only to be once again denied, in Kant's position on mental disturbances, in the *Lectures on Anthropology*. For, so Kant, the demented believes he can attain the absolute totality, inasmuch as "The demented goes beyond the senses (…), demented is then he which substitutes the things of imagination as real." (1997, 108) That is, "they believe they feel in things more than what is really there." (Kant 1997, 108) This happens, according to Kant, because the demented does not envisage the ideal in abstracto, that is, regulatively, "as a means of appreciation (principium dijudicandi)." (1997, 108) Instead, he considers it constitutively, "as the object of desire that we

Of this sort are, for instance, aesthetic ideals.[11]

Therefore, we conclude with Kant, between one infinite and the other infinite – just as between the two lines of human knowledge, and their sciences – there seems to be nothing in common, and these are *rigorously opposed*: "The collective universality, *universitas*, can never be thought when the omnitudo distributiva, *universalitas*, is thought." (1968, 439)

II.2. The Possibility of a *Third Plain of Knowledge*. A Third Infinite, or a Different Manner of Conceiving the Ideal

The previous section – we would say in retrospect – conveyed two important conclusions, which we again present here.

A *first conclusion*, *of a purely metaphysical order,* concerns the infinites' different position in Kant's scheme of human knowledges. Kant's dissociated infinites are different not so much due to their intrinsic characteristics, but especially due to the diverging lines of human knowledge, and the nature of the knowledges themselves, whereupon they labor. For this reason, a Rational Cosmology, a Rational Psychology, an Ontology obey to pure principles and thereby know their qualitative gradation, they (re)cognize themselves intensively and apprehend each other as a totality; whereas an Empirical Psychology, a History, an Empirical Physics depart from the (infinite) cumulation of phenomena and thereby orient themselves quantitatively, (re)cognizing each other by their extension and progressing towards an eternally unattainable maximum.

A *second conclusion*, *of an ontological order,* concerns their actual reach within this scheme. Namely, since the infinites are the consummated image of their lines of knowledge, and since these knowledges are, in both cases, directed at a supreme point, then the infinites are no mere infinites: they rather foreshadow the question of the ideal and are at the basis of all rational and non-rational ideals. A question which is divided between the possibility of attaining the ideal – a finite progression in which the ideal is real –, which Kant does not reject; and the impossibility of attaining the ideal – an infinite progression, through approximation, in which the ideal is ideal – which Kant rejects promptly. Thus could be seen, and thus could be concluded, our presentation of the different infinites in Kant's Metaphysics.

However, it just so happens that precisely here, where it seems to be resolved, *begins the true problem of the ideal* – and this because the problem of the infinites does not end here, in this apparent irreconcilability, just as the problem of Kant's division of human knowledges does not cease here, in the apparent

seek (principium practicum)." (Kant 1997, 108) But, Kant concludes, in this way, that is, "In its most complete degree it [the ideal] is in concreto impossible." (1997, 106)

[11] See Annotations 16 and 17 of this article.

separation between the infinites, and their lines of knowledge. The reason for this assertion demands for prior explanation.

As was seen, it is Kant's view, in the Prolegomena to his *Lectures on Metaphysics*, that the 'ladder' of rational knowledges is opposed, in its finitude, in its distributive character, its systematic nature, to the 'plain soil' of historical sciences, which is infinite, associative and aggregational by nature. And the same happens with the 'reale' and 'mathematicum' infinites, and hence with the two conceptions of ideal that derive from them. But, so suggests Kant, and not by chance precisely on occasion of the exposition of such a fundamental scheme: just because these lines are opposed, and neither can nor should interfere with one another – for they are irremediably different – *this does not mean that they can or should be completely separated*. Quite on the contrary, the lines of human knowledge are not, nor could they be separated – for, according to Kant, "all our knowledge starts with experience (…) but does not come from it." (1968, 615) Instead, *both lines must be united by at least one point*; and if this point cannot be composed by the furthest lengths of each line, which indeed cannot be united, then it should and must be united by the closest point between the two: namely, the point which unites the lowest stadium of the rational line of knowledge to the first one in the horizontal line, and vice-versa. Thus are necessarily united both planes of human knowledge; and thus is constituted, in Kant's perspective, a *perpendicular scheme* in the division of human knowledges, a scheme united by a vertex opening for the communication and connection between rational and empirical sciences, which therefore either permits, or forbids – according to the circumstances – a more or less productive commerce between both.[12]

Now, if this is so, if there can and should be a point of contact, a correspondence between lines and soils of knowledge which nonetheless are opposite, then, since the problem of the fundamental lines of knowledge is indeed the same as that of their infinites, so will this have to apply to this question, and subsequently to that of the ideal. That is, between the reale, rational, finite infinite – the attainable ideal – and the mathematicum, empirical, infinite infinite – the eternally unattainable ideal – *there must also be a relation, perhaps a field of interconnection or mutual dialogue*. Surely not in such a way that the real infinite acts upon the mathematical one, which would be the same as confounding finite and infinite or mistaking the finitude of rational knowledge for the infinitude of the empirical. Nor in such a way that the mathematical infinite claims to be the real infinite, which would be the same as claiming that, according to Kant, the infinite approximation is something to be sustained by reason, and hence, for

[12] A fact which is proved not only by the scheme's own geometrical disposition (a 'ladder' as based upon a 'soil'), but rather theoretically, insofar as some of the lower sciences of applied Metaphysics, such as Rational Psychology and Rational Physics, or Somatology (that is, Rational Cosmology) have a direct correspondence in the horizontal line of knowledge, namely in Empirical Psychology and Empirical Physics (that is, Empirical Cosmology); which means that both must be united by more than just a mere geometrical, formal parallelism.

reason itself, the ideal of reason would be something chimerical. No. The two infinites are indeed different, their knowledges are different, and cannot be united as such. But perhaps one should think that, given the impossibility of merging these two manners of conceiving the ideal, there should emerge *a third one between the latter*: a manner of conceiving the ideal which is neither one, nor the other, yet bears something of both, thereby enhancing what is good in them, their potentialities towards the progression of human knowledge, and at the same time suppressing what is restrictive and negative in both, the limitations they impose to human knowledge. Namely, perhaps it is possible to conceive, in light of Kant's dividing scheme of human knowledges, *a different way to conceive the ideal, one which may operate in a soil which possesses such hybrid characteristics, therefore harboring the vertex between the rational and the empirical and opening for an unsuspected yet very fertile practical field of dialogue between the opposing parties*. In a word, we add, a soil where Kant definitively rethinks the question of the ideal and may set into motion not just another way of conceiving the infinite, or the ideal, but another manner of human knowing and feeling.

For this reason, we inaugurate the third and last section of this article with two guiding questions: 1) Even if we assume that just as there are two modes of cognizing, there are two manners of perceiving the infinite, and that the latter, just as the former, are rigorously opposed – something which Kant unequivocally says – does Kant's work enable us to think a third manner of conceiving the infinite, perhaps one between the infinitum reale and the infinitum mathematicum? And 2) if this proposal is thinkable, in which special soil does Kant see a field for the reception of this third manner of conceiving the ideal, as well as for a new manner of knowing and feeling the I in the World?

III. The Problem of the Ideal from an Anthropological Point of View

III. 1. The Different Ideals. The Singularity of the *Practical Ideal*

As was said in Section II of this text, the other field where Kant approaches the problem of the ideal is that of Anthropology. The problem is raised twice in the *Lectures on Anthropology* (in Kant 1997, 99-100 and Kant 1997, 324-326) and although very brief, and quite similar to that of the *Lectures on Metaphysics*, such a collocation deserves our attention precisely due to that which in it, despite apparently insignificant, is to be distinguished from the others.

According to Kant, the ideal "is the first and most perfect image, according to which all things are possible, or it is an idea in concreto." (1997, 99, 325) And this "idea in concreto" – or "maximum in concreto" (Kant 1997, 325) can be of a triple kind: 1. An "aesthetic ideal"; 2. An "intellectual ideal" (or "ideal of speculative

reason") (Kant 1997, 325); or, at last, 3. A "practical ideal."[13] As if the triple designation was not in itself brief and insufficient, and its theoretical surrounding scarce, Kant approaches only the aesthetic ideal – and briefly at that – considering that this ideal is impossible insofar as *no ideal can be formed merely by sensations.*[14] As to the intellectual-rational and the practical ideals, Kant leaves them unapproached.

In spite of Kant's vague words, it is possible to draw some conclusions on this; especially conclusions which in part confirm the disposition of Section II of this text. For, indeed, Kant refers here first and foremost to two ideals, or two manners of dealing with the problem between finite and infinite. They are, on the one hand, the 'intellectual ideal,' the course of which is the totality of the vertical line of human knowledge and, as such, to paraphrase Kant, is the most perfect image according to which all rational things, thoughts and cognitions are possible – in a word, the most perfect image of a possible, attainable 'infinitum reale.' And, on the other hand, the 'aesthetic ideal,' which is the maximum of the horizontal line of knowledge, and, without surprise, is the most perfect image according to which all aesthetic things and cognitions are possible – the most perfect image of an impossible, unattainable 'infinitum mathematicum.' The first one, as we know, is possible for a number of reasons, already adduced. The second one, Kant again denies; now for several other reasons, which, we think, are worth mentioning. Firstly, because the aesthetic ideal is naturally linked with Aesthetics, and Aesthetics, at least in the vision of the pre-critical Kant – as is here the case, in Collins (1772/73) and Parow (1772/73) – cannot be considered, nor could ever become, a science.[15] Secondly, because since Aesthetics was not a science, and had nothing to do with the laws of reason, it had to be included among the empirical knowledges, along the extensive, infinite line of human knowledge: namely, it constituted a corpus of knowledge which better suited an historical science, and as such had its place among the latter. Finally, if Aesthetics is an historical science,

[13] "We can have three kinds of ideals: 1.) Aesthetic, 2.) Intellectual, 3.) Practical." (Kant 1997, 99) Also: "The ideal is either that from speculative reason, or the aesthetic, or the practical ideal." (Kant 1997, 325)

[14] "As far as the aesthetic [ideal] is concerned, let it be noted that it is not possible to create something from sensations, and, as such, to have an ideal from sensations" (Kant 1997, 99); "In what concerns the aesthetic ideal, no ideal can be formed only from sensations, insofar as that which is said from other-worldly blissfulness are words to which we lack the *concretum*." (Kant 1997, 325)

[15] "The Germans are the only ones who now employ the word 'aesthetics' to designate that which others call the critique of taste. The ground for this is a failed hope, held by the excellent analyst Baumgarten, of bringing the critical judging of the beautiful under principles of reason, and elevating its rules to a science. But this effort is futile. For the putative rules or criteria are merely empirical as far as their sources are concerned, and can therefore never serve as *a priori* rules according to which our judgments of taste must be directed, rather the latter constitutes the genuine touchstone of the correctness of the former." (Kant 1911, 50; CPJ: xix). The same is reiterated throughout the text of *Logik-Jäsche*, "Immanuel Kant's Logik. Ein handbuch zu Vorlesungen." (Kant 1923b, 1-87)

then the 'aesthetic ideal,' that is, the most perfect image according to which all aesthetic objects, all empirical feelings and intuitions, are possible, must simply obey the disposition of the line which embodies it: namely, it must be of an impossible attainment, or to be attainable only through an infinite approximation to the ideal, which for Kant, as is known, is simply unthinkable. In a word, we could then conclude that the intellectual ideal is here evoked in a way reminiscent of that in the *Lectures on Metaphysics*: by confirming its reach, and possibility, under the form of an idea in concreto; as is the aesthetic ideal: as an example of an infinitum mathematicum, as an idea in abstracto, which, just as in Empirical Psychology, in Empirical Physics, in Natural History, is in infinite retraction, as well as in infinite separation, due to its inexhaustible character.

All that is left in the equation is the 'practical ideal' – which Kant does not contemplate in the *Lectures on Metaphysics*, and of which, in truth, we know nothing apart from the fact that is seems to be somehow connected to Kant's anthropological reflection. Yet, by departing from the abovementioned belief that this ideal is no amalgamation of the others, rather an alternative to their rigidity, we propose to resume the guiding vision of our problem – that of the perpendicular scheme of human knowledges – and therein think the possibility of a third dimension of knowledge, as well as *a third ideal*; or, in Kant's words, *a practical ideal*. For, upon doing so, we are imbued with at least one certainty: that, just as the intellectual ideal is the perfect image of a whole set of intellectual objects, and the aesthetic ideal is the perfect image of a set of aesthetic objects, then the practical ideal is to ensure the existence and possible conceivability of a whole series of practical objects – between pure theory and sheer experience – of which the practical ideal must be a perfect image: a truly auspicious belief which we intend to ascertain as well as possible.

III.2. Pragmatic Anthropology, the Abode of the Practical Ideal

In the 1760s and 70s, Kant devotes much care and thought to the question of the mutual reference between the fields of human knowledge – namely, *the task of an encyclopedism of the human spirit* – a problem all the more complex, because not just one, or some, but all domains of human thought demand complete positional and objectual reciprocity amongst themselves.

One of the sciences which, we think, was at the very core of this difficult concatenation of knowledges – for its centrality, for its amplitude, but especially for its unique singularity among all knowledges – is that of *Anthropology*. The reason for this assertion, as well as for these epithets, is quite simple. Anthropology, if not by its attributes as a study of the human being, or by its intimate tone and apparently well-known task,[16] position and destination, seems

[16] "The fault resides certainly in the difficulty of placing this kind of observations, as well as in the singular illusion according to which one believes to know that with which we are accustomed to coexist." (Kant 1997, 7)

to be either nowhere, or everywhere – which is why until the 1770s it had not yet been elevated to science.[17] For, if seen regarding the vertical line of human thought, and its metaphysical sciences, Anthropology, which is indeed anchored upon the empirical and the physiological, seems to have no relation whatsoever with, nor to produce anything of value for the latter; and hence, it seems to have no reference to an 'infinitum reale,' or an intellectual ideal. And yet, quite on the contrary, if seen among the historical sciences, Anthropology seems to pass off as all of these – to be, or to be present in History, Empirical Physics, Empirical Psychology, Medicine, Aesthetics, and hence merely to incur in an infinite model of knowledge, and an infinite ideal: something which the history of modern Anthropology itself proves and ultimately resulted in a rarefaction of the concept of Anthropology, as was often denounced by Kant himself.[18]

Among the many points in Kant's work where the Anthropology is stated in the abovementioned diffuse condition, and where Kant therefore attempts to correct this, by dissociating it from, and reconnecting it to other fields of knowledge, we could stress several. One of them, however, is essential for our line of thought, because it arises precisely in Kant's *Lectures on Metaphysics*, and deals precisely with the position of Anthropology in the perpendicular scheme of human knowledges. For, upon reading these Lectures, one would say that Kant places the Anthropology among the historical sciences; that Kant dissociates Anthropology from rationality and voids it of all contact with reason; in a word, that Kant's voice joins the choir of those which, before him and in his time, considered that Anthropology, its cognitions, its progress as a science, were fixed in the extensive, horizontal line of human knowledge. This because not only here, but in other Lectures on Metaphysics, and other texts, *Kant associates Anthropology and Empirical Psychology, and as such considers it to be an historical science.*[19] And since, according to Kant, Empirical Psychology is merely empirical;[20] since it

[17] Kant refers to this often, specifically in Kant 1997, 7, by asking "Why has a coherent science not derived from the great stock of observations of English authors?" and in Kant 1997, 859, by stating that, by then, "no other book on Anthropology exist[ed]."

[18] Such ambiguity is rendered patent by the very titles of anthropology, or anthropology-like manuals of the time, wherein Anthropology is still very much intertwined with Medicine, Psychology or Physiology, among others: namely, Otto Casmann's *Psychologia Anthropologica* (1594); Riolan's *Anthropographia et Osteologia* (1618); Meisner's *Anthropologia Sacra* (1619); Kyper's *Anthropologia Corporis Humanum* (1647); Sperling's *Synopsys Anthropologiae Physicae* (1659); Hartmann's *Anthropologiae Physico-Medico-Anatomicae* (1696); Teichmeyer's *Elementa Anthropologiae Sive Theoria Corporis Humani* (1719); or even Platner's *Anthropologie für Ärtzte und Weltweisen* (1772), which was of great influence for Kant.

[19] See, regarding Kant's *Lectures on Metaphysics*, Kant 1968, 367; Kant 1980, 757; Kant 1980, 11; see also "Immanuel Kant's Logik. Ein handbuch zu Vorlesungen." (Kant 1923b, 1-87)

[20] "The genuine empirical doctrine of the body most certainly does not belong in Metaphysics; for I cannot speak of air, etc., in metaphysics, for the latter requires empirical principles and therefore belongs to Empirical Physics. The same must I say from the doctrine of the soul; so is Psychology, which serves itself from empirical principles, no pure philosophy." (Kant 1968, 367)

possesses a psychologizing character, just as others have a physiological, or medicinal or anatomical character; since it was once erroneously inscribed in rational soil (in Metaphysics), but does not belong there;[21] since, as such, it cannot be *paired with reason*[22] – that is, it cannot have an express connection with reason, nor, from the reunion of phenomena from experience, provide reason with empirical principles; since it cannot convey a complete knowledge of the human soul, for it has not yet grown to that condition, and, as a result of all the previous, since it cannot be elevated to the condition of philosophical science, nor to that of an academic discipline: [23] then, no doubt, these arguments against Empirical Psychology, which Kant again and again reiterates in the *Lectures on Metaphysics*, must also be the less than favorable attributes of Anthropology, which means that Anthropology is chained to the sciences of erudition and their essential characteristics – the lesser of which is not, to be sure, that of an ideal whose attainment requires an infinite progression, or an infinite approximation – and is impossible. Such must be also, so it seems, Kant's conception of an anthropological (or anthropologically conceived) ideal.

It just so happens, however, that this view of Kant's Anthropology, which seems infallibly right, is taken down in one blow once we compare *Empirical Psychology as Anthropology* – so would Kant often present Anthropology – and *Pragmatic Anthropology,* which emerges in the form of lectures in 1772, but arises much earlier in Kant's thought. For, let it be noted, Empirical Psychology is entirely Anthropology; but Anthropology is not Empirical Psychology except in one of its parts; a part which, already in the text *Logik-Jäsche*, is by Kant surrounded of a Natural Logic and a Natural Aesthetics: namely, Natural Logic=theoretical part of the Anthropology; Natural Aesthetics=Practical part of the Anthropology.[24] And why is this so? Because Kant does indeed refer doubly to Anthropology; and if on the one had Kant does so under the form of an Empirical Psychology, thereby seeming to reduce the boundaries and the scope of Anthropology merely because this is confined to an historical science, on the other hand, when Anthropology appears tacitly (as in the *Lectures on Metaphysics*, or *Logic*) or expressly (as in the *Lectures on Anthropology*) dissociated from Empirical Psychology, then Anthropology, here understood in a pragmatic focus, engulfs Empirical Psychology and has due relations not only with Moral and Religion, but also with Metaphysics.[25] Precisely this says Kant when he refers that Empirical Psychology

[21] On the incorrect inclusion of Anthropology – or even Empirical Psychology – in Metaphysics, see the *Lectures on Anthropology*, Kant 1997, 7-8; Kant 1997, 243; Kant 1997, 473; as well as the *Lectures on Metaphysics*, Kant 1968, 175; Kant 1968, 541; Kant 1980, 750; Kant 1968, 367.
[22] See Kant 1980, 757.
[23] "Psychology has not yet grown to the point that it may convey sufficient data for the knowledge of the soul, so that from it one may create a separate collegium." (Kant 1968, 367)
[24] See Kant 1923b, 16-17.
[25] "The field of Philosophy in sensu cosmopolitico is to be reconducted to the following questions: 1) What can I know? This much is shown by Metaphysics. 2) What should I do? This much is shown by Moral. 3) What should I expect? This much is taught by Religion. 4) What is

was always unduly placed under the wing of Metaphysics, and that as such it is "Metabasis eis allo genos,"[26] whereas (Pragmatic) Anthropology has valid and mutually profitable relations with Metaphysics. As such, we could say, thus answering previous accusations, Pragmatic Anthropology is not merely empirical, rather proceeds by gathering phenomena which it then conveys under the form of principles to Metaphysics – and, therefore, *Anthropology can indeed be paired with reason.*[27] Pragmatic Anthropology bears no foreign character but its own, which characterizes it as the science for the observation of the natural, of the practical, of Man's prudent application in the World – in a word, of the *pragmatic in the human being*. Furthermore, Pragmatic Anthropology has a scope of action and observation much broader than that of Empirical Psychology, to the extent that, *if any science can provide a full knowledge of the human soul, it is Pragmatic Anthropology*; lastly, due to all these reasons, not only is Pragmatic Anthropology, unlike Empirical Psychology, *a science in its own right*, but it could and should be instituted as *an academic discipline* – which it was, precisely by Kant, since 1772. And hence, one may conclude that not only the capacities, the scopes, the positions, the boundaries, the very scientific reach itself, of Empirical Psychology and Pragmatic Anthropology are not the same, but, what is here key, that Empirical Psychology as Anthropology and Anthropology as Pragmatic Anthropology are not the same and must be rigorously distinguished: "of this [Empirical Psychology] may be still distinguished an Anthropology, if by this one understands a knowledge of the human being insofar as it is pragmatic." (Kant 1980, 757)

Now, we ask: what is the result of this *singularization of Pragmatic Anthropology*? This, we think, bears two important repercussions on Kant's scheme of human knowledges, and ultimately on Kant's discussion of the ideals.

The first repercussion has to do with *the position of Pragmatic Anthropology in this scheme*. Namely, if Anthropology, taken in a pragmatic focus, is dissociated from Metaphysics but maintains with it valid relations, and if it is dissociated from historical sciences but still preserves its informative basis in the latter, then this means that Pragmatic Anthropology is due a truly singular, truly hybrid and, at the same time, unique position in the scheme of human knowledges. For, to be sure, Anthropology, which has no real place in any of the lines, but is in both latent, is reserved a place *between* the rational and empirical lines: *a third dimension of human knowledge*, set precisely in the curvilinear space, in the concave wing that opens between the latter. In a word, Pragmatic Anthropology departs from the vertex of the scheme, and engulfs the whole of the angle that opens between

Man? This much is taught by Anthropology. One could deem all of the latter Anthropology, insofar as the first three questions refer to the last one." (Kant 1968, 533-534)

[26] A "μετάβασις είς άλλο γενος," that is, loosely translated, a *complete subversion of the genus of Metaphysics*. See on this Kant 1968, 367; Kant 1980, 757.

[27] According to Kant, Pragmatic Anthropology searches for the "first grounds of the possibility of modification of human nature in general" (1922, 145) and therefore conveys "the subjective principles for all sciences." (1997, 734-735)

intellectual and empirical sciences; but, since it is not meant to be a fusion between the two, then it must acquire the singular, intermediate, mediating characteristics of a whole new knowledge: it, as was said, sets in contact rational laws and empirical cognitions, thus providing these with a possible general validity, and at the same time promoting the reapplication of such empirical data, now under the form of rational laws, upon experience.

The second, final repercussion, even more pungent in our eyes, has to do with *the manner of knowing thus promoted by Pragmatic Anthropology* – which once and for all brings us back to the question of the ideal. For, if before, as he seemed to include Anthropology among the historical sciences and promptly dissociate it from Metaphysics, Kant was in fact collocating part of the Anthropology under the limitations of a line of knowledge which is extensive, which proceeds through addition and is infinite, and separating it from a line of knowledge which proceeds through subsumption, which is intensive and finite, now, however, upon singularizing Pragmatic Anthropology between these two lines, what Kant proposes is *a whole new position for Anthropology*, and hence a *whole new manner of knowing*, hitherto unknown, and yet certainly useful to the latter. What this means, then, is that Pragmatic Anthropology does not obtain its knowledge neither through rational subsumption, nor through empirical addition; its line of knowledge, which is not even a line, rather *a whole dimension*, does not work neither intensively, nor extensively; and as such *its progress towards a totality, or a maximum of cognitions is neither finite, nor infinite*. That is, in Pragmatic Anthropology the question is not that of a finite or an infinite, and much less of an ideal which is restricted to a real or a mathematical infinite: and this because, from the midst of its hybridity, Anthropology not only deals with an indeed apparently inexhaustible collection of data on the human being, and is therefore unquenchable, but it proposes to work towards principles, which are to be general laws of reason, and is therefore quenchable. And hence, what this means is that Anthropology, upon being placed between these more extreme lines of knowledge, thereby occupying an intermediary and mediating function between the two, ensures that this very characteristic and differentiating function is applied to, or embodies, its manner of knowing and hence its knowledge. That is, *its conception of totality, or maximum – its ideal – cannot be but one that is also intermediary, to be found between said real and/or mathematical infinites, and hence between the rational and/or aesthetic ideals*. Namely, Kant proposes here an ideal that, unlike the one which is immediately cognizable and attainable, and unlike the one that is simply incognizable and therefore forever unattainable, *is in fact attainable, yet only through a long progression: one could say, a quasi-infinite – yet indeed finite – duration or progression, thus depriving the rational ideal of what in it is not an ideal, or is insufficiently ideal – its more than possible attainment –, and at the same time depriving the aesthetic-empirical ideal of what is most genuinely ideal in it: its impossible attainment*. This leads us to the conclusion, then, that the ideal proposed by a Pragmatic Anthropology is indeed one consonant

with its dimension in the scheme of knowledges, and with what such a place instils in its manner of knowing: it is a non-rational, non-empirical, rather *a practical ideal*, that is, an ideal which, through a quasi-infinite, and yet also pseudo-finite progression, offers a knowledge that is always an empirical-rational reapplication of itself. To know the natural, or the practical, in Man, is precisely this: not a rational, nor an empirical knowledge of Man's position in the World, rather an extremely long and arduous, yet in the end attainable knowledge which seeks a possible consonance between the rational and the empirical, or, if one prefers, which studies the laws of Nature and their perception by, or integration in, Man's intellect. In a word, the study of the intellect of the I (*Ich als Seele*) within Nature, or the World (*Ich als Mensch*): a "superior anthropological perspective" (Kant 1923a, 374) of Man and the World, or a cosmopolitical doctrine of the World (*Weltlehre*) in the strict sense of the word.

This "superior anthropological perspective" (Kant 1923a, 374) which Kant always connects directly with the possible, yet long and arduous attainment of a practical ideal of knowledge or action, is indeed only alluded to in the *Lectures on Metaphysics*. However, it is not as singular, nor as rare, that it cannot be felt elsewhere. Quite on the contrary, we believe, it arises expressly throughout Kant's whole work, in *Universal Natural History and Theory of the Heavens* (1755) in the *Lectures on Anthropology* (1772-1796), in the texts *Idea for a Universal History* (1784), *On the Common Saying: that may be correct in theory, but it is of no use in practice* (1793), *Toward a Perpetual Peace* (1795) and in *Anthropology from a Pragmatic Point of View* (1798) – therefore summoning, and congregating around its intermediary position, its mediating scope, its finite-infinite manner of knowing, in a word, its extremely long yet finite search for an ideal of humanity, fields as apparently distant as those of Education, Morals, History, Politics or Cosmology.

To convey but a few examples of this quest for a long, yet finite practical ideal, and therewith conclude our line of thought, we summon Kant's words.

As such, already in *Universal Natural History* (1755), the young Kant elevates Man, and the human species, to a teleological understanding of the systematic constitution of the World between Nature and God. Here, Man, as a privileged species, may aspire to "discover the systematic which connects the great members of creation in the great extension of infinity" (Kant 1910, 221) and, by (teleoformically) peering into the secret internal mechanism of the latter, work towards its full formation, and subsequent final perfection. But, so Kant, the attainment of such a practical ideal, which is indeed attainable, may require "millions and mountains of millions of centuries until always new worlds and orders of worlds, ones after the others, may come to form and reach perfection." (Kant 1910, 313)

This "superior perspective" once again emerges in a lecture on Anthropology, *Anthropologie-Friedländer* (1775),[28] a sure precursor of the text *Idea for a Universal History* (1784), where Kant again raises the problem of the necessary full formation of Man's natural dispositions, as, so the title of the lecture, "the character of humanity in general." Here, it is Kant's view that for the attainment of such a desideratum, or a practical ideal of the "supreme degree of perfection" (1997, 697) of the human species, "thousands of years are still required." (1997, 696) Yet, Kant adduces in conformity with said ideal, this practical ideal, though extremely remote and hardly obtainable, "is possible." (1997, 696)[29]

The "superior anthropological perspective" again resurfaces in the text *Idea for a Universal History* (1784). Here, faced with the design of the foundation of a universal history in a cosmopolitical purpose, and the difficulty of its natural consideration merely as "a romance," (Kant 1923a, 29) Kant shows how Man can elevate himself to the condition through work (namely, the anthropological, or practical study and comprehension of his position and teleoformic formation in the World) and, as a species, be himself an integrating and central part in this systematic scheme of things; even though – Kant adds – such a practical ideal may require "a perhaps unending series of generations, each transmitting the other its enlightenment, so as to finally impel [Nature's] germs in our genus towards that degree of development which is completely suitable to its purpose." (Kant 1923a, 19)

And, lastly, this 'superior perspective' emerges in both *Toward Perpetual Peace* (1795) and *On the Common Saying: that may be correct in theory, but it is of no use in practice* (1793), as the superior stage of the comprehension of a relation between the We (or the I as species) and the World: namely, that this superior perspective is superior, and alternative and singular not because it is merely the result of human efforts, or the mere result of an ordination of Nature; that is, to put it in Kantian terms, neither as a merely pragmatic perspective, which inquires on what "as a free-acting being, Man makes of himself, and can and should make [of himself]," (Kant 1917, 119) nor as a merely physiological perspective, which approaches "what Nature makes of Man." (id.; Kant 1923a, 310) Instead, this superior anthropo-cosmological perspective commands us to ponder not what Man makes of himself, nor what Nature makes of Man, *but that which Man makes of himself in consonance with, or along with, what Nature makes of him;*[30] a

[28] "Vom Charackter der Menschheit überhaupt." (Kant 1997, 675-697)

[29] "This consideration is very agreeable, insofar as it is an idea that is possible, for which, however, thousands of years are still required. Nature will always be sufficient, until one such Paradise emerges on Earth. Just as Nature has always formed itself and still does (…), so does the human species, and precisely those many years [140000 years] may pass until the supreme degree of perfection is attained." (Kant 1997, 696-697)

[30] "It [a constant progression towards what is better] depends not so much upon what we do (…), or the method according to which we proceed to bring it about, rather upon what Nature

perspective, an "immeasurably distant success," (Kant 1923a, 310) an ideal to which Man is indeed to elevate himself, and to which Man, once imbued of such a practical disposition, is to access, regardless of its appearance as an ideal between duty and happiness, or a union between politics or morals, or a perpetual peace, or God's kingdom on Earth. A proof, in our opinion, that all these more or less anthropological fields were indeed intimately interconnected in Kant's superior cosmopolitical vision of them; and that, as such, the practical ideal of one is in fact the practical ideal of all others, and this because, in Kant's spirit, the attainment of one was ultimately dependent on the attainment of all others, and vice versa.

References

Allison, Henry E. 2004. *Kant's Transcendental Idealism: An Interpretation and Defense*; Revised and Enlarged Edition. New Heaven and London: Yale University Press.

Erhard, Johann Benjamin. 1970. *Über das Recht des Volks zu einer Revolution und andere Schriften*. München: Carl Hanser Verlag.

Frank, Manfred. 1998. *›Unendliche Annäherung‹. Die Anfänge der philosophischen Frühromantik.* Frankfurt am Main: Suhrkamp.

Franz, Michael. 1996. *Schellings Tübinger Platon-Studien*. Göttingen: Vandenhoeck & Ruprecht.

Gerwen, Rob van. 2009. "Kant's Regulative Principle of Aesthetic Excellence: The Ideal Aesthetic Experience." *Kant-Studien* 86 (3): 331-345.

Haller, Albrecht von. 1882. "[Unvollkommenes Gedicht] über die Ewigkeit." In *Versuch schweizerischer Gedichte*, 149-154. Frauenfeld: J. Huber.

Hölderlin, Johann Friedrich. 1966-1969. *Sämtliche Werke. Grosse Stuttgarter Ausgabe*. Edited by Friedrich Beissner. Stuttgart: W. Kohlhammer Verlag.

Kant, Immanuel. 1901ff. *Gesammelte Schriften*. Hrsg. Von der Königlich-Preussischen Akademie der Wissenschaften zu Berlin (Akademie-Ausgabe). Berlin: Georg Reimer.

————. 1910. *Vorkritische Schriften I*. *Gesammelte Schriften Bd. I* (Akademie-Ausgabe). Berlin: Georg Reimer.

————. 1911. *Kritik der Reinen Vernunft*. *Gesammelte Schriften Bd. III* (Akademie-Ausgabe). Berlin: Georg Reimer.

————. 1913. *Kritik der Praktischen Vernunft. Kritik der Urtheilskraft*. *Gesammelte Schriften Bd. V* (Akademie-Ausgabe). Berlin: Georg Reimer.

————. 1917. *Der Streit der Fakultäten. Anthropologie in Pragmatischer Hinsicht*. *Gesammelte Schriften Bd. VII* (Akademie-Ausgabe). Berlin: Georg Reimer.

makes in us and with us so as to force us to a track to which we would not easily comply by ourselves". (Kant 1923a, 310) A sentence whose content is replicated throughout *Towards Perpetual Peace*, namely, in the section "Of the guarantee of perpetual peace." (Kant 1923a, 360-368)

————. 1922. *Briefwechsel I*. *Gesammelte Schriften Bd. X* (Akademie-Ausgabe). Berlin: Georg Reimer.

————. 1923a. *Abhandlungen nach 1781*. *Gesammelte Schriften Bd. VIII* (Akademie-Ausgabe). Berlin: Georg Reimer.

————. 1923b. *Logik. Physische Geographie. Pädagogik*. *Gesammelte Schriften Bd. IX* (Akademie-Ausgabe). Berlin: Georg Reimer.

————. 1968. *Vorlesungen über Metaphysik und Rationaltheologie*. *Gesammelte Schriften Bd. XVIII* (Akademie-Ausgabe). Berlin: Georg Reimer.

————. 1980. *Kleinere Vorlesungen und Ergänzungen*. *Gesammelte Schriften Bd. XXIX* (Akademie-Ausgabe). Berlin: Georg Reimer.

————.2000. *Critique of the Power of Judgment*. Translated by Paul Guyer and Eric Matthews. Cambridge: Cambridge University Press.

Kleingeld, Pauline. 2012. *Kant and Cosmopolitanism. The Philosophical Ideal of World Citizenship*. Cambridge: Cambridge University Press.

Novalis. 1999. *Schriften. Werke, Tagebücher und Briefe Friedrich von Hardenbergs.* Edited by Hans-Joachim Mähl and Richard Samuel. Darmstadt: Wissenschaftliche Buchgesellschaft.

Schneider, Ruben. 2019. "Kant and the Infinity of Reason." In *The Infinity of God: New Perspectives in Theology and Philosophy*, edited by Benedikt Paul Göcke and Christian Tapp, 78-96. Indiana: University of Notre Dame Press.

Zöller, Günter. 1991. "Kant's Aesthetic Idealism." *The Iowa Review* 21 (2): 52-59.

Information about Authors

Robert James M. Boyles is an associate professor in the Department of Philosophy at De La Salle University, Manila, Philippines. His research interests include, but are not limited to, artificial intelligence ethics, machine ethics, and the philosophy of artificial intelligence. Contact: robert.boyles@dlsu.edu.ph.

Arnold Cusmariu received his Ph.D. in philosophy from Brown University, has taught philosophy at universities in the United States, and has published extensively in metaphysics, epistemology and more recently aesthetics. Publications are available at https://philpapers.org/s/Arnold%20Cusmariu. Also an artist, Dr. Cusmariu received his training in stone carving and bronze casting at the Art League School in Alexandria, Virginia. A sculpture in a series titled *Counterpoint* currently at #24 is pictured in this article. His book *Logic for Kids* is forthcoming from Jenny Stanford Publishing and a book on Picasso's *Demoiselles d'Avignon* is in preparation. Contact: bravo323@gmail.com.

Marlon Jesspher De Vera is Senior Lecturer at the University of the Philippines – Los Baños, Department of Humanities, Philosophy Division, and Senior Lecturer at the University of the Philippines – Diliman, Department of Chemical Engineering. He is also a corporate professional (consumer goods supply chain, purchasing, and manufacturing management) and a doctoral student in the field of Communication at the University of the Philippines Open University. His areas of interest in philosophical research are social and political philosophy, ethics, and metaphysics. Contact: mbdevera2@up.edu.ph.

Landon Frim is an associate professor of philosophy & religious studies at Florida Gulf Coast University. His primary research concerns the intersections of religion, ethics and politics. In this, he investigates the legacy of radical, cosmopolitan thought – from the Stoics and Spinoza, to Marx, Engels, and beyond. His popular work can be found in *Jacobin*, *Salvage Magazine*, *Inside Higher Ed*, and *The New Republic*. Landon's recent book, *Prometheus and Gaia: Technology, Ecology, and Anti-Humanism* criticizes some of the most popular responses towards climate change, and points toward a humanist alternative. His work can be found at https://landonfrim.wordpress.com/. Contact: dfrim@fgcu.edu.

Min Seong Kim is a lecturer in the Graduate Program in Cultural Studies at Sanata Dharma University, Yogyakarta, Indonesia. He is also the chief editor of *Retorik*, an interdisciplinary humanities journal published by the same university. His most recent works include a paper on Ernesto Laclau's discursive ontology in *Philosophia Philippines* and a co-authored chapter on the 'pluriverse' of the

Anthropocene in the edited book *Environmental Governance in Indonesia* (Springer, forthcoming November 2023). Contact: minseong.kim@outlook.com.

Fernando M. F. Silva is a Post-doctoral fellow and member of the Centre for Philosophy at the University of Lisbon. He received his PhD in 2016, on Novalis' critique of identity. He is co-editor of the journal *Estudos Kantianos* and co-coordinator of the Study Nucleus "Kant and German Idealism", at the CFUL. His areas of research are Kantian Aesthetics and Anthropology, German Idealism and Romanticism, with a focus on authors such as Fichte, Novalis or Hölderlin. His main publications include: "'The Poem of the Understanding Is Philosophy'. Novalis and the Art of Self-critique," in *Mimesis Verlag*, Germany (in press). Contact: fmfsilva@yahoo.com.

About the Journal

Symposion was published for the first time in 2003, as *Symposion – Revistă de ştiinţe socio-umane (Symposion – A Journal of Humanities)*, with the purpose of providing a supportive space for academic communication, dialog, and debate, both intra and interdisciplinary, for philosophical humanities and social and political sciences. *Symposion: Theoretical and Applied Inquiries in Philosophy and Social Sciences* shares essentially the same purpose. Its main aim is to promote and provide open access to peer-reviewed, high quality contributions (articles, discussion notes, review essays or book reviews) in philosophy, other humanities disciplines, and social and political sciences connected with philosophy.

The old series published mainly Romanian papers. The new *Symposion* is an international journal, welcoming contributions from around the world written in English and French.

Although devoted especially to social philosophy and related disciplines (such as normative political theory, ethics, social epistemology, philosophy of culture, philosophy of technology, philosophy of economics, philosophy of education, and philosophy of law), the journal is open for original and innovative contributions in all philosophical fields and on all philosophy related topics from other humanities and social and political sciences. *Symposion* is also available for scholars developing interdisciplinary research, as long as it is philosophy related and/or it can open new approaches, pathways, or perspectives for (theoretical or applied) philosophical problems and philosophical thinking. Proposals for special issues devoted to a particular topic, theory or thinker are expressly invited.

The journal promotes all methods and traditions of philosophical analysis and inquiry (from ideal to critical or other types of non-ideal theory, from philosophical hermeneutics to logical and mathematical investigations of philosophical problems, from conceptual analysis to experimental philosophy, and from analytic to Continental philosophy). We also welcome papers on feminist philosophical (and philosophy related) topics, approaches or methods of inquiry.

From 2017, *Symposion* is published on a biannual basis, appearing at the end of May and November. It is published and financed by the "Gheorghe Zane" Institute for Economic and Social Research of The Romanian Academy, Iasi Branch. The publication is free of any fees or charges.

Author Guidelines

1. Accepted Submissions

The journal accepts for publication articles, discussion notes, review essays and book reviews.

Please submit your manuscripts and your proposals for special issues electronically at: symposion.journal@yahoo.com. Authors will receive an e-mail confirming the submission. All subsequent correspondence with the authors will be carried via e-mail. When a paper is co-written, only one author should be identified as the corresponding author.

There are no submission fees or page charges for our journal.

2. Publication Ethics

The journal accepts for publication papers submitted exclusively to *Symposion* and not published, in whole or substantial part, elsewhere. The submitted papers should be the author's own work. All (and only) persons who have a reasonable claim to authorship must be named as co-authors.

The papers suspected of plagiarism, self-plagiarism, redundant publications, unwarranted ('honorary') authorship, unwarranted citations, omitting relevant citations, citing sources that were not read, participation in citation groups (and/or other forms of scholarly misconduct) or the papers containing racist and sexist (or any other kind of offensive, abusive, defamatory, obscene or fraudulent) opinions will be rejected. The authors will be informed about the reasons of the rejection. The editors of *Symposion* reserve the right to take any other legitimate sanctions against the authors proven of scholarly misconduct (such as refusing all future submissions belonging to these authors).

3. Paper Size

The articles should normally not exceed 12000 words in length, including footnotes and references. Articles exceeding 12000 words will be accepted only occasionally and upon a reasonable justification from their authors. The discussion notes and review essays must be no longer than 6000 words and the book reviews must not exceed 4000 words, including footnotes and references. The editors reserve the right to ask the authors to shorten their texts when necessary.

4. Manuscript Format

Manuscripts should be formatted in Rich Text Format file (*rtf) or Microsoft Word document (*docx) and must be double-spaced, including quotes and footnotes, in

12 point Times New Roman font. Where manuscripts contain special symbols, characters and diagrams, the authors are advised to also submit their paper in PDF format. Each page must be numbered and footnotes should be numbered consecutively in the main body of the text and appear at footer of page. Authors should use the author-date system for text citations and chicago style format for reference lists, as it is presented in *Chicago Manual of Style*.For details, please visit http://library.williams.edu/citing/styles/chicago2.php. Large quotations should be set off clearly, by indenting the left margin of the manuscript or by using a smaller font size. Double quotation marks should be used for direct quotations and single quotation marks should be used for quotations within quotations and for words or phrases used in a special sense.

5. Official Languages

The official languages of the journal are English and French. Authors who submit papers not written in their native language are advised to have the article checked for style and grammar by a native speaker. Articles which are not linguistically acceptable may be rejected.

6. Abstract

All submitted articles must have a short abstract not exceeding 200 words in English and 3 to 6 keywords. The abstract must not contain any undefined abbreviations or unspecified references. Authors are asked to compile their manuscripts in the following order: title; abstract; keywords; main text; appendices (as appropriate); references.

7. Author's CV

A short CV including the author`s affiliation and professional postal and email address must be sent in a separate file. All special acknowledgements on behalf of the authors must not appear in the submitted text and should be sent in the separate file. When the manuscript is accepted for publication in the journal, the special acknowledgement will be included in a footnote on the last page of the paper.

8. Review Process

Symposion publishes standard submissions and invited papers. With the exception of invited contributions, all articles which pass the editorial review, will be subject to a strict double anonymous-review process. Therefore the authors should avoid in their manuscripts any mention to their previous work or use an impersonal or neutral form when referring to it.

The submissions will be sent to at least two reviewers recognized as experts in their topics. The editors will take the necessary measures to assure that no conflict of interest is involved in the review process.

The review process is intended to be as quick as possible and to take no more than three months. Authors not receiving any answer during the mentioned period are kindly asked to get in contact with the editors. Processing of papers in French may take longer.

The authors will be notified by the editors via e-mail about the acceptance or rejection of their papers.

9. Acceptance of the Papers

The editorial committee has the final decision on the acceptance of the papers. Articles accepted will be published, as far as possible, in the order in which they are received and will appear in the journal in the alphabetical order of their authors.

The editors reserve their right to ask the authors to revise their papers and the right to require reformatting of accepted manuscripts if they do not meet the norms of the journal.

10. Responsibilities

Authors bear full responsibility for the contents of their own contributions. The opinions expressed in the texts published do not necessarily express the views of the editors. It is the responsibility of the author to obtain written permission for quotations from unpublished material, or for all quotations that exceed the limits provided in the copyright regulations.

11. Checking Proofs

Authors should retain a copy of their paper against which to check proofs. The final proofs will be sent to the corresponding author in PDF format. The author must send an answer within 3 working days. Only minor corrections are accepted and should be sent in a separate file as an e-mail attachment.

12. Reviews

Authors who wish to have their books reviewed in the journal should send them at the following address: *Symposion Journal*, Institutul de Cercetări Economice și Sociale „Gh. Zane" Academia Română, Filiala Iași, Str. Teodor Codrescu, Nr. 2, 700481, Iași, România. The authors of the books are asked to give a valid e-mail address where they will be notified concerning the publishing of a review of their book in our journal. The editors do not guarantee that all the books sent will be reviewed in the journal. The books sent for reviews will not be returned.

13. Copyright & Publishing Rights

their papers but they are kindly asked to inform the Editorial Board of their intention and to provide acknowledgement of the original publication in *Logos & Episteme*, including the title of the article, the journal name, volume, issue number, page number and year of publication. All articles are free for anybody to read and download. They can also be distributed, copied and transmitted on the web, but only for non-commercial purposes, and provided that the journal copyright is acknowledged.

No manuscripts will be returned to their authors. The journal does not pay royalties.

14. Electronic Archives

The journal is archived on theRomanian Academy, Iasi Branch web site. The electronic archives of *Symposion* are also freely available on Philosophy Documentation Center, PhilPapers, Academia.edu, and CiteFactor.