

Erasmus

Erasmus Student Journal of Philosophy

#19 | January 2021

Erasmus Student Journal of Philosophy

www.eur.nl/esphil/esjp
Contact: esjp@esphil.eur.nl

ISSN: 2212-9677

Editorial

As students of Philosophy, we receive great training in creating fascinating lines of reasoning. We become very good at voicing our own standpoint and battling those of others, that is, being at the giving end of the philosophical conversation. Yet, it seems as if we are all a little bit better at giving than taking. Without anybody listening, the purpose of raising one's voice becomes obsolete. Respect and appreciation for another person's opinion is just as important as their freedom to voice it. Reading the works of our fellow colleagues and classmates gives us the unique chance to step into their shoes and carefully follow their line of reasoning. While it is great orators whose names sparkle on the covers of the books in the philosophical canon, it is the listeners who nourish the flame of philosophy throughout the ages.

With this in mind, I would like to present to you five papers written by Rodrigo Bustamante, Matheus Paim, Ina Jüntgen, Matteo Boccacci, and Ahmed AlJuhany. Their articles provide you with an opportunity to practice the art of listening and we are convinced that these authors are worth listening to. Although it is almost impossible to discern the best papers out of many nominations of considerable high quality, the texts presented here are extraordinarily well-written and have left the editorial board truly impressed. We would like to congratulate the authors for their publication. Your hard work, your openness to criticism, and your patience have made your publication in this journal possible. We hope that this experience has carried you further in your process of realizing your great potential as philosophical thinkers. In addition to the papers, this issue also contains an interview with Sonia de Jager about the relation between Philosophy and hands-on engagement in real life.

Bringing out this issue would have never been possible without such a great team of editors, so it is time to give some credits! There have been a lot of changes in the editorial board this year. Many students have joined our team which now consists of 19 editors, including myself as the new editor-in-chief. Also new in his role, though not new to the journal, is our secretary Maximilian Gasser, who has done an amazing job at helping to prepare and coordinate the entire editorial process from beginning to end. Besides the challenges that come with organizing such a large group of "greenhorns", the lingering world pandemic continues to force us into solitude, which is why many members of the editorial board have only met each other as faces on a computer screen. Considering this, it is even more remarkable how well we managed to work together. Under encouragement of our lead editors Matteo Boccacci, Diego Gomez, Lara-Rose Eikamp, Ties van Gemert, and Stijn Voogt, we managed to incorporate the additional obligations for the journal into a busy student schedule. We learned a lot from each other and can be proud of the result of our efforts.

A special thank you goes to the previous editor-in chief Sven Hoogervorst, and to David van Putten. They held our annual editor-training and spent an afternoon of potential free time sharing their wisdom with us. We would also like to acknowledge the amazing advice and support we received from Jamie van der Klaauw and Thijs Heijmeskamp as members of the advisory board, as well as Hub Zwart, Constanze Binder, and Han van Ruler as members of the supervisory board.

We proudly present to you the 19th edition of the ESJP. We would like to invite you to read the papers written by your fellow classmates, your co-authors, and your students. We hope that you will enjoy listening to these writers, that their words will inspire you, that you will carry them into your own discussions, and thereby continue the philosophical conversation.

On behalf of the ESJP editorial board,

Nathalie Maria Kirch
Editor-in-chief

About the Erasmus Student Journal of Philosophy

The Erasmus Student Journal of Philosophy (ESJP) is a double-blind peer-reviewed student journal that publishes the best philosophical papers written by students from the Erasmus School of Philosophy, Erasmus University Rotterdam and from the Humanities Programme of the Erasmus University College. Its aims are to further enrich the philosophical environment in which Rotterdam's philosophy students develop their thinking and bring their best work to the attention of a wider intellectual audience. A new issue of the ESJP appears on our website every January and June.

To ensure the highest possible quality, the ESJP only accepts papers that (a) have been written for a course that is part of the Erasmus University College or Erasmus School of Philosophy curriculum and (b) nominated for publication in the ESJP by the teacher of that course. Each paper that is published in the ESJP is subjected to a double-blind peer review process in which at least one other teacher and two student editors act as referees.

The ESJP encourages students to keep in mind the possibility of publishing their course papers in our journal, and to write papers that appeal to a wider intellectual audience.

Editorial Board

Nathalie Maria Kirch (Editor-in-chief)
Maximilian Gasser (Secretary)
Matteo Boccacci
Gideon Frey
Ina Jäntgen
Pedro Pinilla Plaza
Stijn Voogt
Toine van Mourik
Lara- Rose Eikamp
Diego Gomez
Kas Molenaar
Aycan Funda
Ties van Gemert
Chiara Stenico
Jurian van der Ree
Kas Avedissian
Margot van Baarsen
Ivar Frisch

Advisory Board

Thijs Heijmeskamp
Jamie van der Klaauw

Supervisory Board

Prof. dr. (Hub) H.A.E. Zwart
Dr. (Constanze) C.B. Binder

Prof. dr. (Han) J.A. van Ruler

Disclaimer

Although the editors of the Erasmus Student Journal of Philosophy have taken the utmost care in reviewing the papers in this issue, we cannot exclude the possibility that they contain inaccuracies or violate the proper use of academic referencing or copyright in general. The responsibility for these matters therefore remains with the authors of these papers and third parties that choose to make use of them entirely. In no event can the editorial board of the Erasmus Student Journal of Philosophy or the Faculty of Philosophy of the Erasmus University Rotterdam be held accountable for the contents of these papers.



All work in this issue of the Erasmus Student Journal of Philosophy is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License. For more information, visit <http://creativecommons.org/licenses/by-nc/3.0/>

In this Issue

If you have to decide whether or not to be subject to a medical procedure, you should be able to take the decision autonomously. In “Why Tell Them How: Rethinking Autonomy and Informed Consent in Healthcare”, Ahmed AlJuhany analyzes the complex topic of informed consent in relation to autonomous decisions in healthcare. Do you think it is a good idea to always be aware of what the medical procedure you will be subject to entails? Well, if you think so, you are in good company. In a paper called ‘Autonomy and Consent’, Beauchamp argues that informed consent is a necessary step to foster autonomous decisions. However, Ahmed AlJuhany might surprise you. With a detailed examination of Beauchamp’s argument, the article provides the reader with a thoughtful discussion of the notion of informed consent and of the assumption for which being aware of the basic facts of a medical procedure would always foster autonomy. Ahmed AlJuhany argues that informed consent might harm our capacity to make autonomous decisions. And the paper goes even one step further. AlJuhany argues that invalidating this assumption threatens the entirety of Beauchamp’s account, thus calling for a re-discussion of the notion of informed consent.

How should one go about arriving at a decision when there is a lot of uncertainty? Will a treatment cure all the patients, half of them or none? Ina Jüntgen’s paper is a response to Rowe and Voorhoeve’s thoughts on the decision-making process. Jüntgen specifically addresses situations where there is a lot of uncertainty and the welfare of multiple people is at stake. Jüntgen takes the reader through an example of severe uncertainty and explains concepts that affect our decision-making process, most notably uncertainty aversion and inequality aversion. She concludes that only when both concepts are in conflict, the decision-making process should commence without letting uncertainty aversion have an effect.

In “A Review of Hindriks’ ‘Duty to Join Forces’: A Reducible or Irreducible Collective Responsibility?”, Matteo Boccacci argues that despite the ‘duty to join forces’ leads to an irreducible group obligation, it does not lead to an irreducible accountability. Centered around the issue of whether groups can be considered moral agents, and therefore, have moral obligations, Boccacci focuses on Hindriks’ solution on refusing the agency thesis. However, after presenting this argument, Boccacci moves to present two critical positions which problematize it. The first, a group can still not be held accountable for collective action in the aggregate, but instead, the action is still individually performed. The second, it is still unclear why the sheer presence of other agents would alter individual obligation to count as a group one. To tease this out, Boccacci evaluates three possible scenarios and shows how, in each scenario, the group fails to comply with a “group obligation”. Boccacci then moves to show how each individual may be held accountable, but the entire group could not. Boccacci ultimately argues that in order for Hindriks to make his argument more appealing, he must address the necessity of irreducible obligation to the collective and address the question of how, if accountability still lies at the individual level, there is an irreducible group obligation.

Over the course of 2019, Amsterdam-based high school Cornelius Haga Lyceum was discredited multiple times due to commotion about mismanagement and conflicts of interest within the school’s board, and the school principal’s alleged relations with terrorist groups. On top of several other incidents, the strict islamic orientation of the school caused a debate on the relation between the state and its task to secure education free from interference and the freedom of religious education. In “Promoting freedom of religious

education”, Rodrigo Bustamante effectively breaks down the dilemma at hand. Aiming for a framework that resolves the tension within the debate about religious education, Bustamante presents two different theories of freedom. With the first theory based on positive and negative freedom (Berlin) and a second theory based on autonomy (Christman), Bustamante derives an interesting theory on state’s responsibility to deal with religious education that aligns with the principle of liberal neutrality.

In “Split Mind in a Split Time”, Matheus Paim investigates the role of time-consciousness in the symptomatology and constitution of schizophrenia. Drawing extensively on phenomenological research, he argues that the symptomatology of schizophrenia, including hallucinations, delusions, and disorganized speech, is deeply intertwined with an experience of a disruption of time. After an intricate analysis of Husserl’s philosophy of time, the author suggests that our time-consciousness consisting of retention, primal impression, and protection constitute our experience of enduring objects as well as our awareness of an I stable over time. In schizophrenia, this time-consciousness becomes fragmented. The inner experience of temporal succession no longer corresponds with the duration of the transcendent object and longitudinal intentionality is disrupted. The dissociation that human beings diagnosed with schizophrenia experience between their self and reality results in a disintegration of absolute consciousness and pre-reflective self-awareness. On the basis of this, Paim argues that the structure constituting self and time is similar, almost identical. This intriguing work aims to deepen our understanding of schizophrenia. By seeing how the constitution of self and time are closely related, Paim helps us to uncover what people suffering from this terrible condition undergo.

In addition to the papers, this issue also contains an interview with Sonia de Jager, a second-year PhD student at the Erasmus School of Philosophy. Under the title “Everything is philosophy!”, Ivar Frisch and Ties van Gemert discuss with her whether Philosophy should not be merely about forming objective concepts (of truth, for example). Instead, it should be about hands-on engagement with real life problems and other scientific disciplines. In this engagement, Philosophy should, in a Hegelian fashion, continually reflect on its concept and its goals.

Table of contents

ARTICLES

How we ought to choose when uncertainty aversion conflicts with avoiding inequality **9 - 18**
Amending egalitarianism under severe uncertainty
Ina Jantgen

Why Tell Them How: Rethinking Autonomy and Informed Consent in Healthcare **19 - 27**
Ahmed AlJuhani

A Review of Hindriks' "Duty to Join Forces" **28 - 37**
A Reducible or Irreducible Collective Responsibility?
Matteo Boccacci

Promoting freedom of religious education **38 - 48**
Rodrigo Bustamante

Split Mind in a Split Time. The Disturbance of Internal Time Consciousness in Schizophrenia **49 - 58**
Matheus Paim

INTERVIEWS

"Everything is philosophy!": An Interview with Sonia de Jager. **59 - 67**
Ties van Gemert en Ivar Frisch

How we ought to choose when uncertainty aversion conflicts with avoiding inequality

Amending egalitarianism under severe uncertainty

Ina Jantgen

Section 1: Introduction

Many of us have to make decisions impacting other people's lives daily. And yet, we are often severely uncertain about the consequences our actions will have. In February 2020, for instance, governments around the globe were facing a severely uncertain situation when deciding on shutting down their countries. Lacking sufficient evidence about the novel Covid-19 virus, policymakers could not tell how likely it would be that this virus would kill many people or almost nobody if left to spread uncontrolled. And still, they had to decide how to proceed, impacting many people's lives tremendously. But, when faced with such severe uncertainty, how ought we choose in deciding on the fate of others?

In "Egalitarianism under Severe Uncertainty" (Rowe and Voorhoeve 2018) Thomas Rowe and Alex Voorhoeve propose an answer. They extend a pluralist egalitarian theory of distributive justice to such decisions under severe uncertainty. According to their pluralist egalitarianism, we are morally required to improve people's prospects and how well they end up (Rowe and Voorhoeve 2018, 243f). We are also morally required to avoid inequality in people's prospects and in how well they end up (inequality aversion).

Rowe and Voorhoeve combine this theory of distributive justice with an attitude many people show when faced with severe uncertainty (Rowe and Voorhoeve 2018, 247ff): when unable to assess how likely the possible consequences of their actions are, many people are cautious and give more weight in their decisions to the worst-case scenario than to more favourable ones. According to Rowe and Voorhoeve, these people are uncertainty averse. For the authors, such uncertainty aversion is not just an attitude many people show. It is also rational and morally permissible, including for people making decisions on the fate of others. Thus, Rowe and Voorhoeve address the moral requirements a pluralist egalitarian theory imposes on uncertainty averse people when they make decisions impacting other people's lives.

In this paper, I focus on a class of particularly problematic situations that such an uncertainty averse person can face. Sometimes, the egalitarian concern to avoid inequality conflicts with the concern to avoid uncertainty (Rowe and Voorhoeve 2018, 260f): an action that avoids inequality in how well people end up comes with severe uncertainty about whether those people end up well. Here, Rowe's and Voorhoeve's pluralist egalitarianism does not demand to avoid inequality at the cost of uncertainty or to avoid uncertainty at the cost of inequality. The authors only claim that when deciding what to do one ought to consider avoiding inequality and may permissibly consider avoiding uncertainty. Uncertainty averse people may then decide depending on how they trade-off their uncertainty aversion and inequality aversion.

In this paper, I suggest amending the proposed pluralist egalitarianism for situations in which uncertainty aversion conflicts with inequality aversion. As I argue below, when deciding what to do in such cases, a concern to avoid uncertainty should not be considered. Instead, whenever uncertainty aversion conflicts with avoiding inequality, even uncertainty averse people ought to choose as if they were not uncertainty averse. They may then decide depending on how inequality averse they are.

The paper is organized as follows: In Section 2, I will introduce severe uncertainty and uncertainty aversion. In Section 3, I will describe a case in which uncertainty aversion conflicts with avoiding inequality and how we, according to Rowe and Voorhoeve, ought to choose in this case. In Section 4, I will argue

that uncertainty aversion should not be considered when conflicting with inequality aversion. Instead, whenever both concerns conflict, even uncertainty averse people should choose as if they were indifferent to uncertainty. In Section 5, I will conclude my paper.

Section 2: On severe uncertainty and uncertainty aversion

In this section, I first distinguish the severely uncertain situations I am concerned with from other uncertain situations. I then introduce uncertainty aversion.

To start, in many situations, we are uncertain which consequences our actions will have. Take the following example:

RiskyTreatment: You are a doctor. You want to decide which treatment to give to two patients (call them Lea and Felix). They both have a severe infection, which will leave them brain-damaged if not treated successfully. There are two possible treatments for them: A and B. You know that treatment A will cure Lea but not Felix if their infections are caused by a type 1 virus. In case of a type 2 virus, Treatment A will cure Felix but not Lea. Treatment B will cure both of them if it is a type 1 virus but none of them in case of a type 2 virus. Due to resource constraints, you have to give the same treatment to Lea and Felix. Running some blood tests, you find out that their infections are caused by the same virus. However, the blood tests leave open whether Lea and Felix have a virus infection with a type 1 or a type 2 virus. Luckily though, the tests give you enough information to confidently believe that the chance of Lea and Felix having a type 1 virus is 0.8 and of them having a type 2 virus only 0.2. You now have to choose: do you give Lea and Felix treatment A or treatment B?

In a situation like RiskyTreatment, you face what is known as risk (see Rowe and Voorhoeve 2018, 240f): you do not know what the consequence of your possible actions will be. Here, you do not know what will happen to Lea and Felix if you give them treatment A or B. However, you have sufficient evidence to believe how likely each of the different possible outcomes are.¹ For instance, based on the blood tests, you believe that the chance of treatment A curing Lea but not Felix is 0.8.

Unfortunately, we sometimes face even more severe uncertainty than in RiskyTreatment. Consider the following amendment to the above example:

SeverelyUncertainTreatment: You face the same situation as in RiskyTreatment. However, now the blood tests you run give you no information at all on whether Lea and Felix are more likely to have a type 1 or a type 2 virus. You also have never met Lea and Felix before, have no medical record of them and cannot run any other tests. In contrast to before, you now lack any evidence or prior beliefs on which to base your assessment of how likely it is that they have a type 1 or a type 2 virus. You still have to choose: do you give Lea and Felix treatment A or treatment B?

In SeverelyUncertainTreatment, you face severe uncertainty² (Rowe and Voorhoeve 2018, 248; see also Bradley and Drechsler 2014, 1233–37; Ellsberg 1961): based on your evidence and prior beliefs, you cannot tell how likely each of the different possible consequences of your potential actions is. In contrast to before, you only believe that the chance of these consequences occurring is within some range.³ For instance, based on the poor evidence of your blood test, you might believe that the chance of treatment A curing Lea but not Felix is between 0.2 and 0.8. But, in contrast to your previous situation, you lack sufficient evidence to reasonably believe that it is precisely 0.8.

¹Following Rowe and Voorhoeve, I use risk as a subjective concept, that is as referring to the beliefs about how likely some outcome obtains that a rational person can form based on their evidence and prior beliefs (Rowe and Voorhoeve 2018, 241).

²Severe uncertainty about the possible states of the world is also referred to as ambiguity (Bradley and Drechsler 2014).

³I also use severe uncertainty as a subjective concept (Rowe and Voorhoeve 2018, 241; see also footnote 1).

Now, suppose you could choose between facing RiskyTreatment and SeverelyUncertainTreatment. Which situation would you prefer to be in? There is evidence that many people would choose RiskyTreatment (Rowe and Voorhoeve 2018, 247ff; see also Trautmann and Kuilen 2015). These people behave uncertainty aversely.

According to Rowe and Voorhoeve, such uncertainty averse behaviour is motivated by a cautious attitude in light of severe uncertainty (Rowe and Voorhoeve 2018, 248f): uncertainty averse people cautiously give greater weight in their decision to the least favourable probability of an outcome occurring which they deem possible than to more favourable ones. To illustrate, consider treatment B in SeverelyUncertainTreatment. Recall this treatment cures both Lea and Felix if they have a type 1 virus but none of them in case of a type 2 virus. Suppose you believe that the probability that they have a type 1 virus is between 0.2 and 0.8. Then, the least favourable probability you consider possible is 0.2. It is least likely then that treatment B cures both Lea and Felix. The most favourable probability you consider possible is 0.8. It is most likely then that treatment B cures both Lea and Felix. According to Rowe and Voorhoeve, as an uncertainty averse person, you then give the former possibility, 0.2, greater weight in your decision on Lea's and Felix' treatment than more favourable possibilities such as 0.8. In such a way, you express a cautious attitude in light of your severe uncertainty. For Rowe and Voorhoeve, you are uncertainty averse.

Rowe and Voorhoeve maintain that it is rational and morally permissible to be uncertainty averse (Rowe and Voorhoeve 2018, 248–50). For them, however, it is not rationally or morally required (Rowe & Voorhoeve, 2018, p. 243). One could also give the same weight to all probabilities of the consequences of one's actions that one deems possible. For instance, you could give the same weight in your decision to 0.2 and 0.8. You would then not express a cautious attitude when faced with severe uncertainty – you would be uncertainty neutral. You could also give less weight to the least favourable probability that you deem possible than to other more favourable ones. For instance, you could give less weight to 0.2 than to 0.8. You would then express an adventurous attitude towards uncertainty – you would be uncertainty seeking.

Nevertheless, for Rowe and Voorhoeve, uncertainty aversion is a common and permissible attitude. Therefore, they focus on uncertainty averse people (Rowe and Voorhoeve 2018, 250). They discuss how an uncertainty averse person ought to choose when deciding on the fate of others, as in SeverelyUncertainTreatment.

In this paper, I focus on a class of particularly problematic situations uncertainty averse people can face - situations in which avoiding uncertainty comes at the cost of inequality. SeverelyUncertainTreatment is such a case. In the next section, I use this example to describe the conflict between uncertainty aversion and avoiding inequality and how we, according to Rowe's and Voorhoeve's pluralist egalitarianism, ought to choose in this case.

Section 3: When uncertainty aversion conflicts with inequality aversion

Before explaining the conflict between uncertainty aversion and avoiding inequality, I first describe the decision you face in *SeverelyUncertainTreatment* from a pluralist egalitarian perspective.

Section 3.1: SeverelyUncertainTreatment from a pluralist egalitarian perspective

Let us consider the choice you face in *SeverelyUncertainTreatment* from the viewpoint of Rowe’s and Voorhoeve’s pluralist egalitarianism (see Rowe and Voorhoeve 2018). I illustrate your choice in Table 1.

| Possible action | | Type 1 virus | Type 2 virus | Prospective value of individuals’ prospects | Prospective value of possible anonymized distributions of final well-being |
|-----------------|-------|-------------------|-------------------|---|--|
| | | $0 \leq p \leq 1$ | $0 \leq p \leq 1$ | $\alpha = 0.6$ | $\alpha = 0.6$ |
| Treatment A | Lea | 90 | 30 | 54 | $v\{90,30\}$ |
| | Felix | 30 | 90 | 54 | |
| Treatment B | Lea | 90 | 30 | 54 | $0.6 * v\{30,30\} + 0.4 * v\{90,90\}$ |

Table 1: Your choice in SeverelyUncertainTreatment

To start, we can measure how well Lea and Felix would end up if treated successfully or not (see Rowe and Voorhoeve 2018, 244). Recall that, if not treated successfully Lea and Felix end up with brain damage. Their final well-being would be quite low (measured as 30). By contrast, if treated successfully they would be fully cured. Their final well-being would be quite high (measured as 90).

Furthermore, in *SeverelyUncertainTreatment*, you do not assign precise probabilities to Lea and Felix having a type 1 or a type 2 virus. I will assume here that you consider any probability of them having either virus possible. In other words, you believe that the probability of a type 1 virus or of a type 2 virus is respectively between 0 and 1 ($0 \leq p \leq 1$).

Rowe’s and Voorhoeve’s pluralist egalitarian theory then cares about (a) improving people’s prospects, (b) avoiding inequality in people’s prospects, (c) raising total final well-being and (d) avoiding inequality in final well-being (Rowe and Voorhoeve 2018, 243f). To apply this theory to our severely uncertain situation, we have to understand how to evaluate the consequences of our actions for individuals’ prospects and distributions of final well-being in such situations.⁶ In other words, we need to understand how to obtain the prospective value of individuals’ prospects and of the possible distributions of final well-being (Rowe and Voorhoeve 2018, 253). Let us start with Lea’s and Felix’ prospects.

To calculate the prospective value of an individual’s prospects under severe uncertainty, Rowe and Voorhoeve use a particular decision criterion called α -maxmin (Rowe and Voorhoeve 2018, 251).⁷ According to α -maxmin, the prospective value of an individual’s prospect is α times the expected value of the prospect under the worst possible probability distribution and $1-\alpha$ times the expected value of the prospect under the best possible probability distribution.⁸ For *SeverelyUncertainTreatment*, we get: $\alpha * (30 * 1 + 90 * 0) + (1-\alpha) * (90 * 1 + 30 * 0)$.

Which value does α have for an uncertainty averse person? Recall, for Rowe and Voorhoeve, an uncertainty averse person gives more weight to the worst possible probability distribution over possible outcomes of their actions than to more favourable ones (see previous section). In the α -maxmin criterion, this is expressed

⁶We cannot calculate the expected value of Lea’s and Felix’s prospects and of the distribution of final well-being. To do so we would need to assign a precise probability to the different possible final well-being levels and the distributions of final well-being. In a severely uncertain situation, according to Rowe and Voorhoeve, we cannot assign such precise probabilities.

⁷Rowe and Voorhoeve maintain that their results hold for other popular decision criteria one could use (Rowe and Voorhoeve 2018, 250–51).

⁸This implies that the pluralist egalitarian view does not consider only the worst possible scenario.

⁹This number is arbitrary.

as $\alpha > 0.5$ (Rowe and Voorhoeve 2018, 251). For SeverelyUncertainTreatment, using $\alpha = 0.6$, we get the following prospective value of Lea's and Felix' prospects for both treatments: $0.6 \cdot (30 \cdot 1 + 90 \cdot 0) + 0.4 \cdot (90 \cdot 1 + 30 \cdot 0) = 54$.

Next, we need to consider the prospective value of the possible anonymized distributions of final well-being. To understand how this prospective value is calculated, we need to understand how pluralist egalitarianism evaluates distributions of final well-being. Following Rowe and Voorhoeve, I use $v\{x,y\}$ for the value of a distribution in which one person has x final well-being and the other y final well-being (Rowe and Voorhoeve 2018, 257 footnote 28). The authors' pluralist egalitarianism then cares both about avoiding inequality in final well-being and improving final well-being (see Otsuka and Voorhoeve 2018, 81f; and also Voorhoeve and Fleurbaey 2016). Taking these concerns into account, possible distributions of final well-being are evaluated as follows:

An equal distribution of final well-being is evaluated as the average of final well-being in it. For instance, $v\{60,60\} = 60$. The pluralist egalitarian concern to avoid inequality then reduces the value of an unequal distribution of final well-being below the average of final well-being in it. Suppose, for example, that you ought to be indifferent between a distribution of 90 final well-being for one person and 30 final well-being for another person and a distribution of 50 for both persons. This means $v\{90,30\} = 50$ instead of $v\{90,30\} = 60$. Moreover, the pluralist egalitarian concern to improve well-being implies that a distribution with lower average final well-being is less desirable than one with higher average final well-being. This means $v\{50,50\} < v\{60,60\}$. In such a way, the pluralist egalitarian view assigns a value to each possible distribution of final well-being. In SeverelyUncertainTreatment, we get: $v\{90,30\}$, $v\{30,30\}$ and $v\{90,90\}$.

To calculate the prospective value of the possible distributions of final well-being, Rowe and Voorhoeve again use α -maxmin (Rowe and Voorhoeve 2018, 257 footnote 28). The only difference to the case of individuals' prospects is to replace individuals' possible final well-being levels with the possible values of the distributions of final well-being. Here this means:

$$\begin{aligned} \text{prospective value of possible distributions of final well-being in A} &= \alpha \\ &\cdot (1 \cdot v\{90,30\} + 0 \cdot v\{90,30\}) + (1 - \alpha) \cdot (0 \cdot v\{90,30\} + 1 \cdot v\{90,30\}) = v\{90,30\} \\ \text{prospective value of possible distributions of final well-being in B} &= \alpha \\ &\cdot (1 \cdot v\{30,30\} + 0 \cdot v\{90,90\}) + (1 - \alpha) \cdot (0 \cdot v\{30,30\} + 1 \cdot v\{90,90\}) = \alpha \cdot v\{30,30\} + (1 - \alpha) \cdot v\{90,90\} \end{aligned}$$

According to the proposed pluralist egalitarianism, you then ought to choose an action yielding the highest prospective value of possible distributions of final well-being (see Rowe and Voorhoeve 2018; and also Otsuka and Voorhoeve 2018, 81f; Voorhoeve and Fleurbaey 2016, 941f).

Having described your choice in SeverelyUncertainTreatment from a pluralist egalitarian perspective, I turn now to how uncertainty aversion conflicts with inequality aversion in SeverelyUncertainTreatment.

Section 3.2: Pluralist egalitarianism when uncertainty aversion conflicts with inequality aversion

Let us look at what the proposed pluralist egalitarianism requires you to consider when deciding on Lea's and Felix' treatments in SeverelyUncertainTreatment.

Rowe's and Voorhoeve's pluralist egalitarian theory requires you (a) to improve people's prospects and (b) to avoid inequality in people's prospects (Rowe and Voorhoeve 2018, 243f). Here, both treatments give Lea and Felix the same prospects. Thus, neither fares better concerning these.

Moreover, the proposed pluralist egalitarian view requires you to choose the highest prospective value of the

possible distributions of final well-being, expressing a concern to (c) raise final well-being and to (d) avoid inequality in final well-being (see Rowe and Voorhoeve 2018; and also Otsuka and Voorhoeve 2018, 81f; Voorhoeve and Fleurbaey 2016, 941f). Does treatment A or treatment B have a higher prospective value in SeverelyUncertainTreatment? This depends on two concerns: on your uncertainty aversion (i.e. on α) and on inequality aversion (i.e. on how low the value of the unequal distribution, $v\{90,30\}$, is).

In cases like SeverelyUncertainTreatment, these two concerns conflict (Rowe and Voorhoeve 2018, 260f). Treatment A will result in an unequal distribution of final well-being, while treatment B results in an equal distribution. Inequality aversion then speaks against A and in favour of B. But B involves uncertainty about the distribution of final well-being. You do not know how likely it is that you cure both Lea and Felix rather than none of them. By contrast, treatment A leaves you certain that the treatment will cure one person, leaving the other with brain damage. Your uncertainty aversion then speaks against B and in favour of A. In short, uncertainty aversion and inequality aversion suggest opposing actions.

Which action ought you choose if a concern to avoid uncertainty and one to avoid inequality point in opposite directions? Here, the proposed pluralist egalitarianism does not demand to choose inequality at cost of uncertainty or uncertainty at cost of inequality (Rowe and Voorhoeve 2018, 261). Instead, according to Rowe's and Voorhoeve's theory, so long as you also consider to avoid inequality in your decision, your uncertainty aversion can permissibly be a reason in favour of choosing an unequal distribution (Rowe & Voorhoeve, 2018, p. 265). You may then choose either action depending on how inequality and uncertainty averse you are (Rowe and Voorhoeve 2018, 261). If you care a lot about inequality but are only slightly uncertainty averse, B has a higher prospective value for you.¹¹ You choose B. By contrast, if you are very uncertainty averse but only slightly care for inequality, then A has a higher prospective value for you.¹² You choose A. For Rowe and Voorhoeve, both choices are permissible.

In the next section, I propose to amend Rowe's and Voorhoeve's theory for cases in which uncertainty aversion conflicts with avoiding inequality. I argue that when deciding in such cases it is not permissible to consider one's uncertainty aversion as a reason in favour of choosing an unequal distribution. Rather, whenever both concerns conflict, you ought to choose as if you were uncertainty neutral. You may then choose either action depending on how inequality averse you are.

Section 4: Amending pluralist egalitarianism for conflicts between uncertainty aversion and inequality aversion

The amendment I propose to the suggested pluralist egalitarianism relies on a difference in the moral importance of uncertainty and inequality within the view: inequality is per se morally objectionable while uncertainty is a moral burden only for uncertainty averse people. Let me start by explaining this difference. Consider inequality. As outlined above, Rowe and Voorhoeve extend pluralist egalitarianism to severely uncertain situations. This theory views inequality in individuals' prospects and final well-being as morally objectionable (Otsuka and Voorhoeve 2018; Voorhoeve and Fleurbaey 2016). Inequality ought to be avoided by any person.

Contrast this with uncertainty. Within the proposed pluralist egalitarianism under severe uncertainty, uncertainty is not per se morally objectionable such that any person ought to avoid it. This view only considers how uncertainty is a moral burden for an uncertainty averse person (Rowe & Voorhoeve, 2018, p. 252). Yet, as noted earlier, Rowe and Voorhoeve do not argue that uncertainty aversion is a morally required attitude (Rowe & Voorhoeve, 2018, p. 243). You could permissibly be uncertainty neutral or

¹¹This would mean for you: $\alpha * v\{30,30\} + (1-\alpha) * v\{90,90\} > v\{90,30\}$

¹²This would mean for you: $\alpha * v\{30,30\} + (1-\alpha) * v\{90,90\} < v\{90,30\}$

seeking. However, if it is permissible to be uncertainty neutral or seeking within this theory, then this theory does not claim that uncertainty should be a moral burden for any person. Hence, in the proposed theory, uncertainty is not per se morally objectionable – contrary to inequality.

One might object that Rowe's and Voorhoeve's pluralist egalitarianism could still imply that uncertainty is per se morally objectionable. This would be the case if uncertainty would constitute a moral burden for uncertainty neutral or seeking people in a similar way as it does for uncertainty averse people within pluralist egalitarianism.

However, Rowe's and Voorhoeve's pluralist egalitarian theory does not imply that uncertainty is per se morally objectionable. To see this, consider why uncertainty is a moral burden for uncertainty averse people: First, uncertainty reduces the prospective value of individuals' prospects for uncertainty averse people (see Rowe and Voorhoeve 2018, 265). Under uncertainty aversion, the more uncertain you are about individuals' prospects the lower is the prospective value of these prospects for you. The authors' pluralist egalitarian position requires both to improve these prospects and to avoid inequality in them (see previous section). Thus, the reductive effect of uncertainty about individuals' prospects is a moral burden for any uncertainty averse person. In *SeverelyUncertainTreatment*, given you are uncertainty averse, uncertainty reduces the prospective value of Lea's and Felix' prospects. Due to this reductive effect, uncertainty makes treatment A and B morally worse options than in the absence of such uncertainty.

Second, for uncertainty averse people, uncertainty reduces the prospective value of the possible distributions of final well-being. Under uncertainty aversion, the more uncertain you are about these distributions the lower is the prospective value of these distributions. Within pluralist egalitarianism, one ought to choose an action with a higher prospective value of the possible distributions of final well-being (see previous section). Hence, the reductive effect of uncertainty about the distributions of final well-being is a moral burden for an uncertainty averse person. In *SeverelyUncertainTreatment*, given your uncertainty aversion, the uncertainty about the possible distributions of final well-being in B reduces the prospective value of these distributions. Due to this reductive effect, uncertainty makes B a morally worse option than it would be in the absence of such uncertainty.

Now, consider what happens if someone is not uncertainty averse. For an uncertainty neutral or seeking person, uncertainty does not reduce the prospective value of people's prospects or of the possible distributions of final well-being. This is because this person would give equal or less weight to the least favourable possible probability distribution over the consequences of their actions. In lack of a reductive effect on the prospective value of people's prospects or of the distributions of final well-being, uncertainty is not a moral burden within pluralist egalitarianism. Thus, uncertainty is not a moral burden for uncertainty neutral or seeking people. Instead, it is only a moral burden for uncertainty averse people. Given this, uncertainty is not per se morally objectionable – contrary to inequality.

This difference in moral importance between uncertainty and inequality provides a reason to amend Rowe's and Voorhoeve's theory for cases in which uncertainty aversion conflicts with inequality aversion. Instead of allowing decision-makers to trade these concerns off, uncertainty aversion should be ignored whenever conflicting with inequality aversion.

To see why note first how uncertainty averse people consider avoiding morally objectionable inequality less in their decisions than uncertainty neutral or seeking people do. To illustrate, suppose you are very uncertainty averse and only slightly care for avoiding inequality in *SeverelyUncertainTreatment*. Your uncertainty aversion would drastically reduce the prospective value of the distributions of final well-being which treatment B could generate. This value would be very low, making treatment B a morally bad option for you. By contrast, your slight inequality aversion amounts to a relatively high prospective value of the

unequal distribution of final well-being generated by treatment A. Your slight inequality aversion makes A a morally acceptable option for you. The prospective value of A would be higher for you than the one of B. You would choose A – the unequal distribution.

Crucially, in your decision for A, you consider the morally objectionable inequality in A less than an uncertainty neutral or seeking person would. For an uncertainty neutral or seeking person, the prospective value of the possible distributions of final well-being under B would not be as low. B would be a morally better option for this person than it is for you as uncertainty averse. For the same slight inequality aversion, this uncertainty neutral or seeking person would then choose B – the equal distribution. In short, your uncertainty aversion leads you to consider avoiding morally objectionable inequality less in your decision than an uncertainty neutral or seeking person does.

Now, note, secondly, how ignoring your uncertainty aversion leads you to consider the moral burden of uncertainty less in your decision than if you uphold uncertainty aversion. In such a case, the prospective value of the distributions of final well-being under B would be higher for you. B would be a morally better option than it is if you uphold your uncertainty aversion. Even if you are still only slightly averse to inequality, you would now choose B – the equal distribution. Notably, even if you ignore your uncertainty aversion, uncertainty is still a moral burden for you as an uncertainty averse person. Hence, in this case, you consider the moral burden of uncertainty in B less in your decision than if you were upholding your uncertainty aversion.

However, due to the difference in moral importance between inequality and uncertainty, considering the moral burden of uncertainty less in your choice is not en par with considering inequality less. Despite being a moral burden for you, uncertainty is not per se morally objectionable within the proposed pluralist egalitarianism. In contrast to upholding your uncertainty aversion, ignoring your uncertainty aversion then only reduces how much you consider what is a moral burden exclusively for uncertainty averse people rather than morally objectionable per se.

I maintain that from a moral point of view it is more important to fully consider what is per se morally objectionable than to fully consider what is a moral burden only for some people. Therefore, it should not be permissible for your uncertainty aversion to reduce how much you consider avoiding morally objectionable inequality, as Rowe and Voorhoeve claim. Rather, their pluralist egalitarian view should be amended: Upholding one's uncertainty aversion should not be morally permissible whenever it conflicts with avoiding inequality.

How should you then evaluate the possible distributions of final well-being in *SeverelyUncertainTreatment*? It follows that you should evaluate these as if you were uncertainty neutral.¹³ This avoids letting uncertainty aversion reduce how much you consider avoiding morally objectionable inequality.

Let me clarify: I do not propose to amend pluralist egalitarianism under severe uncertainty such as to require people to be uncertainty neutral. Instead, I argue that whenever uncertainty aversion would conflict with inequality aversion, one should decide as if one were uncertainty neutral. This does not imply that one should be uncertainty neutral when both concerns do not conflict (as in most cases Rowe and Voorhoeve discuss).

There is still an open question: what ought you choose in *SeverelyUncertainTreatment*? I have argued that you should be uncertainty neutral. This means you should adopt $\alpha=0.5$ to obtain the prospective values of the possible distributions of final well-being. Here, this means: $0.5 \cdot v\{90,90\} + 0.5 \cdot v\{30,30\}$ for B and $v\{90,30\}$ for A. You then still ought to choose the action with a higher prospective value of the possible distributions of final well-being.

¹³One should not decide as uncertainty seeking (implying $\alpha < 0.5$). This would imply that uncertainty is claimed to be morally desirable, which in the proposed pluralist egalitarian view it is not.

But which is higher? This depends on the values of the distributions of final well-being. This means it depends on how inequality averse you are. For instance, if you are very inequality averse, you choose B, accepting the uncertain chance of leaving both Lea and Felix with brain damage to avoid inequality in final well-being. If you are only slightly inequality averse, you choose A, allowing inequality in final well-being to avoid the uncertain chance of leaving them both with brain damage. I have not argued how you ought to evaluate the distributions of final well-being. Thus, I also leave open whether you ought to choose A or B. I only maintain that you should not give extra weight to the uncertain chance of leaving Lea and Felix both with brain damage. You should decide as if you were uncertainty neutral.

Section 5: Conclusion

In severely uncertain situations, decisions on actions impacting other people's lives are difficult to make. In such situations, we lack sufficient evidence to say precisely how likely different possible consequences of our actions are. In light of this, many people cautiously consider the worst-case scenario more than others in their decisions. According to Rowe and Voorhoeve, they express an uncertainty averse attitude. Rowe and Voorhoeve maintain that such uncertainty aversion is rationally and morally permissible. Their pluralist egalitarianism under severe uncertainty addresses how uncertainty averse people ought to choose when deciding on the fate of others. In some especially problematic cases, uncertainty aversion conflicts with their morally required concern to avoid inequality. Here, the authors do not demand a particular action. Instead, for them, uncertainty averse people may decide depending on how they trade off their concern to avoid uncertainty and the egalitarian concern to avoid inequality.

I have argued that we should amend Rowe's and Voorhoeve's pluralist egalitarianism for situations in which uncertainty aversion and inequality aversion conflict. Within their pluralist egalitarian view, inequality is morally objectionable, while uncertainty is only a moral burden for uncertainty averse people. This difference in moral importance gives a reason not to let uncertainty aversion be considered whenever conflicting with avoiding inequality. Considering uncertainty aversion reduces how much consideration is given to morally objectionable inequality. By contrast, ignoring uncertainty aversion only reduces how much consideration is given to uncertainty. And uncertainty is a moral burden exclusively for uncertainty averse people, rather than per se morally objectionable. Therefore, whenever uncertainty aversion conflicts with inequality aversion, one should ignore one's uncertainty aversion and decide as if one were indifferent to uncertainty.

Bibliography

Al-Najjar, Nabil I., and Jonathan Weinstein. 2009.

“The Ambiguity Aversion Literature: A Critical Assessment.” *Economics & Philosophy* 25 (3): 249–84. <https://doi.org/10.1017/S026626710999023X>.

Bradley, Richard, and Mareile Drechsler. 2014.

“Types of Uncertainty.” *Erkenntnis* 79 (6): 1225–48. <https://doi.org/10.1007/s10670-013-9518-4>.

Ellsberg, Daniel. 1961.

“Risk, Ambiguity, and the Savage Axioms.” *The Quarterly Journal of Economics* 75 (4): 643–69. <https://doi.org/10.2307/1884324>.

Fleurbaey, Marc. 2018.

“Welfare Economics, Risk and Uncertainty.” *Canadian Journal of Economics/Revue Canadienne d'économique* 51 (1): 5–40. <https://doi.org/10.1111/caje.12314>.

Otsuka, Michael, and Alex Voorhoeve. 2018.

“Equality Versus Priority.” In *The Oxford Handbook of Distributive Justice*, edited by Serena Olsaretti. Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199645121.013.1>.

Rowe, Thomas, and Alex Voorhoeve. 2018.

“Egalitarianism under Severe Uncertainty.” *Philosophy & Public Affairs* 46 (3): 239–68. <https://doi.org/10.1111/papa.12121>.

Trautmann, Stefan T., and Gijs van de Kuilen. 2015.

“Ambiguity Attitudes.” In *The Wiley Blackwell Handbook of Judgment and Decision Making*, 89–116. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118468333.ch3>.

Voorhoeve, Alex, and Marc Fleurbaey. 2016.

“Priority or Equality for Possible People?” *Ethics* 126 (4): 929–54. <https://doi.org/10.1086/686000>.

Why Tell Them How: Rethinking Autonomy and Informed Consent in Healthcare

Ahmed AlJuhani

Instilled in almost all our medical institutions and research centers is the requirement for physicians and investigators to obtain informed consent from patients and subjects before carrying out medical interventions. The requirement of informed consent has become so central to modern medical practice that its theoretical justification is rarely questioned by the general public¹. But, within Bioethics, one recurring argument in its defence is that informed consent protects patients' autonomy – their capacity for self-determined action in the absence of controlling influences (Beauchamp and Childress 1985, 102). This is the view that Beauchamp (2010) defends. It is the view that I will argue against.

Beauchamp's defence of this view rests on a unique understanding of autonomy. He develops an account - called the "concise theory" – that takes autonomy to be a matter of degrees. Under the concise theory, actions are not defined in binary terms as either being autonomous or non-autonomous. Instead, they are only more or less autonomous depending on the extent to which they are (i) intentional, (ii) carried out with adequate understanding and (iii) in the absence of controlling influences (*ibid.*, 65). In outlining this theory of autonomy, Beauchamp purports to have presented a clear definition of the term, one that makes explicit the conditions that lead to its realisation. He then moves on to demonstrate how informed consent protects autonomy by ensuring that these conditions are met.

In this article, I will argue against Beauchamp's claim that the requirement of informed consent protects autonomy through these conditions and should be justified on that basis. To do this, I will first show that the process by which we ensure the condition of adequate understanding does not protect patients' autonomy and can, in fact, contradict it. Secondly, I show how the remaining conditions of intentionality and the absence of controlling influences are insufficient in protecting patients' autonomy. Therefore, I conclude that informed consent does not always protect autonomy and should not be justified on the basis that it does.

This essay is structured as follows: in section 1, I review Beauchamp's argument, clarifying his notion of concise autonomy, and I illustrate how informed consent is thought to protect it. In section 2, I present two requirements determining whether (A) the condition of adequate understanding is satisfied and (B) whether the condition succeeds in protecting autonomy. The first requirement specifies what information should be understood for the condition of adequate understanding to be satisfied, while the second requirement specifies when satisfying this condition also protects a patient's capacity for autonomous action. I show that the first requirement can be met in ways that contradict autonomy, while the second cannot be said to have been met. After this is established, I go on to argue that informed consent does not necessarily protect autonomy. In section 3, I recap and conclude.

¹ An illuminating exercise is to review surveys of clinician, patient, and public opinion about informed consent (to start with, see Singer 1993, El-Wakeel et al. 2006, and Riordan et al. 2015). It becomes apparent that few, if indeed any, survey opinions about why informed consent matters in the first place.

1. The Concise Theory of Autonomy and Its Relation to Informed Consent

The concise theory of autonomy takes the term to be non-binary: actions are neither autonomous nor non-autonomous, they are only more or less so. This sliding-scale notion of autonomy depends on the extent to which one's actions are (i) intentional, carried out with (ii) adequate understanding and (iii) in absence of controlling influences (Beauchamp 2010, 65). Beauchamp recognises that, for the theory to be workable, it would require a cut-off point where an action can safely be labelled autonomous or non-autonomous. Beauchamp argues (*ibid.*, 71) that cut-off points should be determined according to context. Under this account, an action would be considered autonomous or non-autonomous in light of "specific objectives of decision making such as deciding about surgery, choosing a university to attend, and hiring a new employee" (*ibid.*, 71).

Before I go on to show how informed consent is thought to protect autonomy through these conditions, we will have to take a closer look at what the concise theory entails. There are two points that are worth clarifying: first, how should we interpret to what each of the conditions laid out above – intentionality, adequate understanding and absence of controlling influence – amounts to? Second, how ought we to interpret the notion of cut-offs that Beauchamp describes? As Beauchamp (*ibid.*, 71) recognises, establishing cut-offs is necessary for a workable notion of autonomy. In other words, it is imperative that we have a clear understanding of where we draw the line between autonomous and non-autonomous action. I will clarify both points below before I go on to elaborate Beauchamp's argument that informed consent protects autonomy.

1.1 Conditions of Autonomy: Intentionality, Adequate Understanding and the Absence of Controlling Influence

Let us begin with the condition of intentionality. Beauchamp describes an action as intentional if and only if it is willed according to a plan (Beauchamp 2010, 66). This does not necessarily imply that every intentional action will lead to a predicted outcome. Beauchamp (*ibid.*, 67) accepts that even intentional action can occasionally tolerate undesirable consequences if they are a part of a more general plan (or, perhaps a tolerable feature of one). For example, if a patient decides to undergo a root canal treatment to relieve herself of recurring pain that is caused by a deep decay in her tooth, then the prospect of experiencing mild pain after a successful procedure is only a tolerable (though unfortunate) byproduct of her plans. If she does in fact suffer some pain after the procedure, then this is simply an undesirable part of her general plan to receive treatment.

Beauchamp's take on adequate understanding is more intricate. In his view, a patient must acquire "pertinent information and have relevant beliefs" (*ibid.*, 68) about the nature and consequences of her actions. It's important to note that neither pertinence nor relevance imply perfection or completeness. A patient can have adequate understanding only by grasping some basic, material facts about the procedure (*ibid.*). If a patient decides to undergo a root canal treatment and authorises the dentist to carry out the procedure, she need only obtain some information about what this procedure entails: for example, whether a dental x-ray is required, whether a local anesthetic will be used, and whether there is a risk of experiencing any pain after the procedure. According to Beauchamp (*ibid.*), knowing these basic facts and being able to form beliefs (or preferences) about them means that the patient's understanding is adequate.

Beauchamp's last condition – the absence of controlling influence – is rather straightforward. It demands that a patient should be free of controls exerted either by external sources or by internal states that rob her of the ability to direct her own actions (ibid., 69). Beauchamp (ibid., 70) considers several cases that border on controlling influences including coercion and manipulation. Coercion is defined as the use or threat of harm to enforce an action, while manipulation is defined as a non-persuasive means of altering a patient's opinion (either by withholding or misrepresenting information). To give an example of both: one can imagine a physician influencing a reluctant patient to undergo a diagnostic exam either by threatening to abandon her if she does not comply (coercion) or by misrepresenting the costs and pains often suffered by those who undergo this examination (manipulation). The absence of such influences is a necessary requirement for protecting a patient's autonomy.

The conditions are clear enough when they are considered separately. But a complication arises when we stop to consider how they may come together to determine autonomy. According to the concise theory, it is the extent to which these conditions are satisfied that determines whether a patient's actions are more or less autonomous. But this raises the obvious question: at what point do we say that the patient's actions are autonomous? When are we justified in defining any particular cut-off point? These are the questions I will now address.

1.2 Locating Cut-off Points in the Concise Theory of Autonomy

How should we understand Beauchamp's recommendation that there ought to be a cut-off point above which the patient's actions are considered autonomous, below which they are not? As described above, Beauchamp (ibid., 71) argues that the cut-off point should be determined by the context. But, on its own, this statement is unclear. There are at least two ways in which we may interpret Beauchamp's position here: either we understand cut-off points as absolutely defined according to specific conditions that patients might find themselves in, or we understand them as relatively defined according to each specific patient's perception of the particular situation in which she finds herself.

Under the first interpretation, the cut-off point is externally assigned and fixed for every possible context – regardless of the acting patient's concerns, her beliefs or preferences². If a patient is undergoing a minor procedure – say, a root canal – then her decision can be considered autonomous only if she meets an externally assigned degree of autonomy. If a patient is undergoing a major procedure – say, an open-heart surgery – then her decision can be considered autonomous only if she meets a different (presumably higher) degree of autonomy. The central point of this interpretation is that the same cut-off point is applied to all patients as long as they are undergoing the same procedure.

But this might not be the most charitable reading of Beauchamp's position. This is for the simple reason that it implies that decision scenarios can be defined absolutely regardless of the patient's concerns and of the contingencies affecting her decisions. This is no minor issue. To maintain an absolute interpretation of cut-off points, we would have to accept either that the patients' concerns are negligible or that the absolute interpretation can encompass every patient's concerns in every possible scenario, and under every possible contingency. The first commitment flies in the face of autonomy since, by definition, it holds that the patients' concerns do not matter.

² By 'externally assigned' I mean to say that the cut-off is defined not by those participating in the procedure, but by any other source - perhaps an arm-chair philosopher, who elaborates some criteria from which we derive the cut-off required.

The second commitment imposes a daunting epistemic requirement for us to define a patient's actions as autonomous or non-autonomous.

To see this, consider how we would establish an absolute cut-off point for root canals. The treatment may be considered a minor procedure by an adult. But if that same patient was younger, then the same procedure might be considered more significant. If a patient is suffering a serious illness, then a root canal treatment's imposition might seem negligible in comparison. If the patient starts to heal, then the root canal's imposition relatively grows in significance. Taking such examples at large scale, it would be difficult to define a cut-off point for every person, every procedure and every possible contingency. There might as well be an infinite number of relevant variables to consider.

A second, more charitable reading of Beauchamp's position would simply state that there is no absolute degree of autonomy that ought to be satisfied. Instead, locating cut-off points turns into a pragmatic exercise. It is a subjective judgement when a patient's actions are considered autonomous or non-autonomous, decided after taking into account our own aims, the patient's concerns and the contingencies affecting them³. Beauchamp (*ibid.*, 71) himself alludes to a similar, pragmatic interpretation when he warns that a theory of autonomy may - in a given context - set too high or low a threshold to be workable. In the former case, we risk rendering "nonautonomous ... [what] many individuals normally regard as autonomous", while in the latter case we risk rendering "many individuals who are normally regarded as non-autonomous [as] autonomous" (*ibid.*). One implication of this reading is that a patient's actions are only sufficiently autonomous according to the context in which they are carried out. A certain degree of autonomy might be sufficient in light of certain conditions, but insufficient in light of others. For example, an adult patient who decides to undergo a root canal treatment with *x* degree of autonomy might be considered sufficiently autonomous, while a minor who decides to undergo the same treatment with the same degree of autonomy might not be. In light of our institutions, the minor would require her legal guardian's consent. Sufficiency can therefore be interpreted as context-specific.

So far, I have reviewed Beauchamp's concise theory of autonomy, clarifying its conditions and its cut-offs. This theory develops an account of degrees of autonomy, which relies on the fulfilment of three conditions: (i) intentionality, (ii) adequate understanding, and (iii) the absence of controlling influences. Recognising that such an account is only workable once we decide when an action can aptly be labelled as autonomous or not, Beauchamp (*ibid.*, 71) recommends that we establish cut-off points in light of our objectives and in light of the contingencies at hand. In the next subsection, I will show how informed consent is thought to protect autonomy.

1.3 How Informed Consent Protects Autonomy

Beauchamp (*ibid.*, 57) tries to demonstrate that informed consent protects autonomy by showing that it ensures the three conditions laid out by the concise theory. He does this by defining informed consent as the autonomous authorisation of a procedure - meaning that, in healthcare, a patient gives informed consent if she authorises a medical intervention (i) intentionally, with (ii) adequate understanding and in (iii) absence of controlling influence. By defining informed consent in terms of autonomous action (authorisation), Beauchamp ensures that by satisfying the former's requirements, we simultaneously satisfy the latter's conditions.

³ Although such an interpretation might resemble the problematic case previously described - where one would have to define an absolute cut-off point for every case - it differs in that it allows us to simply drop certain contingencies from consideration in light of our aims and the patient's concerns. It is pragmatic in that we can define a patient's actions as autonomous or not in considering however many contingencies that we think are useful. It also does not commit us to any strong claim that there is a well-defined degree of autonomy that should be met by all patients. Hence, it is far more lenient than the absolute interpretation of cut-off points.

To see how this is done, let us borrow from a previous example. Suppose that we are worried about the possibility that a domineering physician might threaten to abandon a patient if she does not agree to undergo a diagnostic exam. Assuming that it is obtained without failure, the requirement of informed consent would protect the patient from such coercion by ensuring that she agrees to the examination only in the (iii) absence of controlling influences (ibid., 60). Similarly, suppose we are worried that the same, domineering physician might manipulate her patient by omitting information about the diagnostic exam. Once again, requiring informed consent can alleviate this worry because it demands that the patient agrees to this examination only if she has (ii) adequate understanding (ibid., 60).

Now we can see how Beauchamp's argument fits together – how he develops an account of autonomy as a matter of degrees and how he ensures that they protect autonomy. Having established Beauchamp's position, I now turn to argue against it.

2. Informed Consent Does Not Protect Autonomy

My argument takes three steps to make: first, I will show that for adequate understanding to be satisfied in a way that protects autonomy it will need to satisfy two requirements: (A) determines whether adequate understanding has been satisfied; (B) determines whether adequate understanding protects patients' capacity for autonomous action. Next, I argue that (A) can be met in ways that contradicts autonomy, while (B) cannot be said to have been met. This, I will argue, tells us that adequate understanding is not concerned with protecting autonomy. The third and final step of my argument is to show that the remaining two conditions are insufficient in protecting autonomy.

2.1 Two Additional Requirements for Adequate Understanding

Beauchamp's description of adequate understanding relies on lenient requirements. As mentioned, he argues that a patient satisfies this condition only if she has a basic understanding of "pertinent information" of the procedure to which she subjects herself and can formulate beliefs about them (Beauchamp 2010, 68). But this requirement leaves much to be desired. Which facts are most pertinent? How would one determine them? Since Beauchamp does not give any explicit answers to these questions, we are left to interpret them from the text.

Here we face the same interpretative issue that we faced during our discussion of cut-off points. One may interpret what is pertinent as absolute – externally assigned and fixed for every possible context, irrespective of the patient's identity. Alternatively, one may interpret pertinence as relative – pragmatically assigned according to the patient's identity and circumstance. For the very same reasons that has led us to reject the absolute interpretation of cut-off points, I will reject the absolute interpretation of pertinent information: it would impose the daunting task of defining these terms for every person, every procedure, and every possible contingency.

We are now left with a relativist interpretation of pertinent information. Like the case with cut-offs, the relativist stance is also a pragmatic one, deciding what is pertinent according to the physicians' goals, the patient's identity, and her beliefs and preferences. But notice that this interpretation raises a difficult problem: the information most pertinent to the patient may not be the information that the physician deems pertinent. This makes it possible for the physician's goals to conflict with the patients' beliefs and preferences.

If we prioritise the patients' interpretation of what is pertinent over the physicians', then we risk allowing for the patient to undergo an invasive procedure without knowing what it entails. To see this, consider again the

example of a patient undergoing a root canal treatment. Suppose that she does not care much about what the treatment actually entails nor about its side effects. Instead, this patient only cares about whether her dentist endorses this procedure and whether this procedure is affordable⁴. All other concerns are secondary and negligible.

The problem here is that neither concern tells the patient anything about the procedure itself or about its side-effects. Yet the patient can still satisfy the condition of adequate understanding since she understands the information she finds most pertinent: she is sure that the dentist recommends the procedure and that she can afford it. In this case, satisfying the condition of adequate understanding requires no actual knowledge of what a root canal treatment actually entails⁵.

Of course, this is an intuitively undesirable conclusion. Presumably, an ideal definition of adequate understanding should not allow for such awkward results. We would want for there to be a set of non-arbitrary facts that the patient should know before she undergoes a medical procedure. Such facts should, for example, ensure that the patient has an idea of what a procedure entails, what side-effects she should expect, and whether there are any alternative ways of addressing her medical problem. Hence, for the condition of adequate understanding to be met, the patient would need to satisfy an additional requirement:

- A. For adequate understanding to be achieved, the patient must know a set of non-arbitrary, material facts about the procedure.

But there is also a different requirement to be met. We would want to know if meeting the condition of adequate understanding necessarily protects patients' autonomy. Beauchamp (*ibid.*, 60) himself hints at such a requirement when he argues that adequate understanding protects autonomy when it "help[s] patients ... improve the quality of their decision making, which is a matter of fostering autonomous choice." Hence, our second condition:

- B. For adequate understanding to enhance autonomy, it should be a fact that improving the patients' understanding protects her capacity for autonomous actions.

Here, I conclude the first step of my argument. I have established two additional requirements for determining whether (A) the condition of adequate understanding is met and (B) whether the condition succeeds in protecting autonomy. In the next section, I will move on to the second step of my argument to make the case that both requirements can be met in ways that contradict autonomy. I then proceed to the third step of my argument to demonstrate that the remaining conditions – intentionally and the absence of controlling influences – are insufficient in protecting autonomy.

2.2 Contradicting Autonomy with Adequate Understanding

Having established requirements (A) and (B), let us examine them in detail. For (A) to be realised, one must accept that we care more about what the patient ought to know than what the patient cares to know. This is implied by the fact that requirement (A) demands that patients must be made aware of certain information that we (or, perhaps the physicians) believe to be most pertinent, regardless of what the patient wants or cares to know.

⁴ One might of course argue that the patient wisely trusts the experts' opinion, but this does not change the fact that she might not know what a root canal treatment entails.

⁵ Though this might seem unrealistic, it is worth remembering that we are describing a hypothetical scenario that illuminates the strength and weakness of a particular account of informed consent. If there are such faults in simple cases like these, then we can imagine that similar faults could arise in situations that are more serious and complex.

This ultimately means that, for the physician to satisfy requirement (A) and meet the condition of adequate understanding, he may have to enforce information upon the patient that she does not wish or care to know.

Imagine again that a patient is deciding whether to undergo a root canal treatment. But in this case, imagine that she also has a crippling fear of pain and does not want to hear about the possibility of suffering an infection or an inflammation after the procedure. To satisfy requirement (A) of informed consent, the physician would have to inform the patient about the possible complications that might occur after the procedure has taken place. Otherwise, the patient would not know the whole set of material facts that requirement (A) has deemed necessary. In a case like this, requirement (A) obliges the dentist to contradict the patient's desires and inform her of the procedure's possibly painful side-effects. The dentist might have to infringe on the patients' autonomy by going against the patient's intention to not know certain facts that were nevertheless deemed pertinent. If instead we decide that this is an undesirable conclusion – that we should not abide by requirement (A) after all – then we would revert back to the original problem of arbitrary judgements: the patient would be allowed to be ignorant about the procedure to which she subjects herself.

What about requirement (B)? For the requirement to be realised, one would have to show that by improving a patient's understanding of a procedure we protect her capacity for autonomous action. But it is unclear whether informing the patient about pertinent facts does any such thing. What I want to do here is reject Beauchamp's (*ibid.*, 6.0) answer that it necessarily does. Imagine, the same pain-fearing patient that is about to consent to a root canal treatment. Suppose that her dentist somehow determines that only 10% of people suffer mild pains as a result of minor complications after the procedure, whereas 1% suffer more serious pains as a result of major complications.

It is possible that by informing the patient of this, the dentist is actually hindering her capacity for autonomous action. In laboratory settings, it's been repeatedly shown that fearful individuals often overestimate the probabilities of harmful outcomes (for a review, see Weimer and Pauli 2016). In our case, it is possible that the patient vastly overestimates the probability of her experiencing serious pain and therefore refuses to undergo the procedure altogether. This would be especially problematic if she would have chosen otherwise had she an accurate estimate of the likelihood of a major complication occurring and of her suffering any serious pain (assuming there is such a thing as an 'accurate estimate'). It's a tricky matter, but we cannot exclude the possibility that the dentist might unintentionally drive the patient to act in a way that she would not want. In this case, the dentist may have imposed an external, controlling influence, thereby contradicting one of the conditions of concise autonomy⁶. We therefore cannot accept Beauchamp's claim that meeting adequate understanding necessarily protects the patient's capacity for autonomous action, at least not without further qualification⁷.

What lessons can we draw from this discussion? From the first one, we learn that satisfying requirement (A) has less to do with protecting autonomy and more to do with the patient being informed about the procedure she considers. This means that our concern with ensuring the patient's understanding overrides our concern for the patient's capacity as an autonomous actor. Our conclusion here casts doubt over adequate understanding's status as a condition for autonomy.

⁶ Note, that this does not at all imply that a patient who overestimates harmful probabilities cannot ever act autonomously. All this means is that we cannot grant Beauchamp's claim that more information inevitably improves the patient's capacity for good and autonomous action. In other words: more information does not necessarily protect autonomy.

⁷ Pursuing this line of reasoning would lead us too far afield. For our purposes, it is sufficient to say that we cannot accept Beauchamp's claim about adequate understanding without further qualification.

This concludes the second step in my argument: to show that adequate understanding does not necessarily protect autonomy and can, instead, contradict it. Now, I move on to the third and final step of the argument: I show that the remaining conditions of intentionality and the absence of controlling influences are insufficient in protecting autonomy.

2.3. Why Informed Consent Does Not Protect Autonomy

Recall that informed consent is defined as an autonomous authorisation of a procedure. If the preceding argument is right, adequate understanding cannot be said to protect autonomy and so is not implied by the word 'autonomous' here. The only remaining conditions of autonomy are intentionality and the absence of controlling influences. I will now construct an example where informed consent – understood as intentional authorisation, absent of controlling influences – does not protect autonomy and yet is still desirably obtained. Imagine the following scenario:

A physician administers for a recovering alcoholic a pain management program after she had undergone major surgery⁸. This pain management program includes a potentially addictive opioid medication. Suppose the patient is unaware that the medication may prove to be addictive or euphoric enough to tempt a relapse (since adequate understanding is not meant to protect autonomy, it needn't feature in our example). In fact, she has a provisional desire not to know of any negative side-effects that might deter her from taking the pain-relieving medication. She then takes the medication intentionally and in the absence of controlling influences. In a few weeks time, the patient relapses and once again submits to her addictive tendencies.

Let us take a closer look at this scenario. First, the recovering alcoholic – being sober – is fully capable of acting intentionally. At the time, she can decide whether she should accept the medication and can act on her decision according to a plan that she wills. By taking the medication, risking a relapse and finally relapsing, we can say that the patient's intentional actions today have stripped her of the capacity for intentional action tomorrow. Ensuring that she acts intentionally has, in this case, failed to protect the patient's autonomy.

What about the absence of controlling influences? One might be tempted to argue that the patient was an alcoholic and was therefore influenced by some psychological disposition that was beyond her control. But this would be wrong. In this scenario, we imagine that the patient is content in recovery and is in a perfect state to make her decision. And, in our example, the patient's desire not to know of any negative side-effects is only the result of her desperate desire to alleviate pain, not a sly, subconscious desire for a euphoric 'fix'.

One might argue that had the patient been adequately informed about the medication's addictive risks, then the patient's autonomy would have been protected. But this is not necessarily the case. First, it would contradict the patient's explicit desire not to know about any side-effects that might deter her from taking the medication and alleviating her pain. This, as we have established earlier, might constitute an infringement of the patient's autonomous intentions. Second, if we suppose that the patient is glad to know of the medication's addictive risks, then this still does not undermine the point that intentionality and an absence of controlling influences might not protect her autonomy.

⁸ To avoid confusion: a recovering alcoholic – as is commonly described – is one who has experienced what it is like to be addicted to alcohol, has suffered its detriments, and is now living soberly. The recovering alcoholic is not influenced by any substance. She merely knows that if she drinks again, she is almost guaranteed to succumb to her addictive tendencies.

The patient may decide that the risk is worthwhile: she intentionally and without controlling influences decides to take the potentially addictive medication. If she relapses, we would not be able to conclude that her autonomy was protected.

Now, if adequate understanding does not protect autonomy, and if the remaining conditions of intentionality and an absence of controlling influence are insufficient for the same purpose, then we are left with a puzzling result. Why would we still require informed consent in situations where Beauchamp's conditions fail to protect patients' autonomy? One possible answer is that informed consent is not about autonomy at all but about trust, dignity or mutual respect. This would fit cases where autonomy is protected and cases where it is not. Insofar as we are looking to justify the requirement of informed consent, we should look to justify it on grounds that are valid in a variety of cases – not only when it serves one particular purpose. Hence, I conclude that informed consent should not be justified on the basis that it protects autonomy.

Conclusion

I have argued against Beauchamp's claim that informed consent protects autonomy and should be justified on that basis. I began by considering Beauchamp's definition of autonomy, which depends on the conditions of intentionality, adequate understanding and the absence of controlling influence. My argument then proceeded in three steps: first, I showed that adequate understanding would need two additional requirements that determine (A) whether it is met and (B) whether it protects autonomy. Second, I argued that both requirements can be met in ways that contradict autonomy. In the third and final step of my argument, I argued that informed consent does not always protect autonomy if its remaining conditions only were intentionality and an absence of controlling influences. I conclude that the requirement of informed consent cannot be justified on grounds that it protects informed consent.

References:

1. Beauchamp, Tom L. 2010. "Autonomy and consent." *The ethics of consent: Theory and practice*. 55-78.
2. Beauchamp, Tom L., and James F. Childress. 2001. *Principles of biomedical ethics*. Oxford University Press.
3. El-Wakeel, H., Gordon J. Taylor, and J. J. T. Tate. 2006. "What do patients really want to know in an informed consent procedure? A questionnaire-based survey of patients in the Bath area, UK." *Journal of medical ethics* 32, no. 10 : 612-616.
4. Riordan, Fiona, Chrysanthi Papoutsis, Julie E. Reed, Cicely Marston, Derek Bell, and Azeem Majeed. 2015. "Patient and public attitudes towards informed consent models and levels of awareness of Electronic Health Records in the UK." *International journal of medical informatics* 84, no. 4: 237-247.
5. Singer, E. 1993. Informed consent and survey response: a summary of the empirical literature. *Journal of Official Statistics*, 9 (2): 361.
6. Wiemer, Julian, and Paul Pauli. 2016. "Fear-relevant illusory correlations in different fears and anxiety disorders: A review of the literature." *Journal of anxiety disorders* 42: 113-128.

A Review of Hindriks' "Duty to Join Forces"

A Reducible or Irreducible Collective Responsibility?

Matteo Boccacci

Introduction

In a recent paper, Hindriks (2019) claims that random groups of individuals can bear irreducible group responsibilities. To clarify the conceptual issue at stake, consider this example from Björnsson (2020):

Offshore wind: Like most summer days, a large group of children are enjoying the beach, playing on their air mattresses close to land. This particular day, there are fifteen of them; three came with a parent, the other twelve live nearby. Without warning, the weak onshore wind quickly turns strongly offshore, and the children begin drifting out, beyond the range of their swimming capabilities. They need to be rescued. In response, each of the three adults can swim out and catch one child before the wind has carried them too far, provided that they start swimming now. But there is also a lifeboat that could be dragged a few yards to the sea if at least two adults joined forces; with it, two of them could row out and pick up all the children. Each of the adults quickly realizes all this.

Offshore wind is similar to the famous drowning child example from Singer (1972). If an individual walking by a shallow pond notices a child drowning in it, she is morally obliged to save the child. In offshore wind, none of the three adults could save the children alone. However, if the adults coordinate their actions, the entire group of children could be saved with a similar ease and serious harm could be prevented. Thus, it seems reasonable to assess that in offshore wind a moral obligation to save the children weighs on the adults (Hindriks 2019, 208; Björnsson 2020, 3)⁹.

In the literature, there is an ongoing discussion to discern whether we should ascribe the obligation to the group of adults or to each adult part of the group. In technical terms, the issue is specified as follows: on the one hand, we could have a reducible group obligation. The group is indeed obliged to save the children, but this obligation that weighs on the group is nothing else than the sum of the obligations that each individual bears. On the other hand, we could have an irreducible group obligation. We consider the group obligation as weighing on the group itself, thus being unequal to the individual obligation that weighs on the individual members.

The pivotal point of the discussion is the agency thesis (AT): generally, it is argued that only moral agents can bear moral obligations (Björnsson 2020; Hindriks 2019).¹⁰ Attempts to argue that it is the group as such that bears the moral obligation have to deal with AT. It is possible to either claim that AT is false, thus claiming that also non-agential groups can bear obligations, or that groups can be taken to be moral agents distinct from the summa of individuals' agencies.

⁹ Note that it is possible to disagree on the existence of the obligation. Williams (1981) argues that it would be 'one thought too many' for the adults to reflect whether to save the entire group or only one child considering that she is their own. I leave this discussion out since it is not fundamental to the aims of my paper. I assume Björnsson's thesis that there is a moral obligation to save the entire group of children. I thank an anonymous editor for making me aware of this point.

¹⁰ A moral agent is an agent that is able to distinguish between good and bad and that can act on those considerations. Consider the case of a newborn: it is not possible to argue that she is a moral agent insofar as she cannot distinguish between good and bad and you would not consider her morally responsible for her actions or for bearing any sort of moral obligations.

With a famous contribution, List and Pettit (2011) have argued that some structured groups (e.g., corporations, states) can qualify as moral agents thus respecting the validity of AT while ascribing to structured groups the possibility to be morally responsible. The argument that List and Pettit (2011) bring out does not apply to offshore wind. In the example, we have a random group of individuals (RGI) (Held 1970). In the literature, RGIs are often taken to fail in meeting the agency requirements (Hindriks 2019; Björnsson 2020; Smiley 2017, 24-25).¹¹ Thus, ascribing to these groups moral responsibility would represent a violation of AT.

Despite this difficulty, different scholars have looked for strategies to claim that RGIs can bear irreducible responsibilities (Björnsson 2014; Hindriks 2019; Wringer 2010). In this paper, I will focus on one solution based upon the refusal of AT. Hindriks (2019) argues that non-agential groups can bear irreducible group responsibility, thus claiming that AT is false. This is possible in virtue of what he calls 'The Duty to Join Forces.'

The remainder of the paper is structured as follows: in section 1, I present Hindriks' (2019) argument. In section 2, I move to a first criticism of his position. I argue that, irrespective of Hindriks' argument, it seems to be the case that the collective action of the group is not accountable, and that accountability still lies on the individual level. In section 3, I move to present a second criticism to Hindriks' argument: his irreducibility seems to lose reliability, thus, questioning his argument.

1. The Duty to Join Forces

Hindriks (2019) claims that we can make the case for irreducible responsibilities for RGIs analyzing what he calls 'The Duty to Join Forces' (DJF). Hindriks describes a process based on two phases: firstly, there is a stage in which individuals approach each other and make each other aware of their willingness to prevent harm. Secondly, the collective brings about a joint action to prevent it.

In offshore wind, first stage, we would have something like this:

Good people in Offshorewind (1/2): One of the three adults asks the other two to collaborate. Suddenly, one of the two answers that she will join helping with the lifeboat and the other one stays aside willing to help if needed.

Since there is a moral obligation to save the entire group of children, Hindriks (p. 210) claims that the three adults have a lateral responsibility to join forces: to coordinate their actions and prevent the harm, the three individuals have a responsibility to group up. In practice, one would have to approach the other(s) and ask for their help to save the drowning children. In this case, the other individual(s) would have an individual moral obligation to agree to help in case their contribution is needed to prevent the harm.

This moral obligation to group up is conditional on the actual possibility for the group to be able to save the children: if, for example, four adults were needed to prevent the harm but only three were present on the beach, the group of adults would be excused to only save their own children. To see it in concrete, consider a variation of Scarcity from Björnsson (2020):

Scarcity: We are in offshore wind, besides that, the boat is bigger and now, to prevent the death of the children, three adults (a-b-c) should coordinate. Adult a starts running to save her child careless of the other children. Adult b asks Adult c if she wants to try to use the boat to save the entire group. Adult c refuses and starts running to save her child.

¹¹ Some structured groups are usually considered to meet the agency requirements because they have a structure and a decision process (List and Pettit 2011). It is more difficult to ascribe agency to RGIs because they lack both these requirements.

According to Hindriks (2019, p. 206-209), adult *c* would not be blameworthy. Assuming that it is evident that two adults cannot handle the bigger boat, adult *c* is justified to refuse to help and save her child. The moral obligation to answer affirmatively when asked to collaborate depends upon the possibility of success of the action. According to Hindriks, the individual is morally obliged to help insofar as the action is likely to succeed.

So far, the three salient aspects of the DJF are the following: a) the three adults have a moral obligation to save the entire group of children. b) The three adults cannot do it on their own. c) Individuals have a lateral responsibility to join forces in order to prevent the harm; this obligation in the first stage relies on the individual level and they are individually accountable for it.

After agreeing on acting together to prevent the collective harm, in the second stage of the DJF, the group performs the joint action and saves the group of children. In offshore wind, we would have something like this:

Good people in Offshorewind (2/2): We are in good people in offshorewind(1/2). The two adults that agreed on acting take the boat and go save the children. They pick-up the entire group of them and they all get back to the beach safely.

If the first and the second stage of the DJF are respected, Hindriks argues that the adults would comply with their moral obligation to save the children. It is not clear yet whether we can consider this obligation reducible to the individual obligations of each member or whether it is an irreducible obligation distinct from them. In the next subsection, I present Hindriks' argument for the irreducibility of the DJF.

1.1 The irreducibility of the Duty to Join Forces

Before presenting Hindriks' argument, it might be useful to spend a few words on reducibility. Hindriks (2019) is not explicit in what kind of reducibility he is referring to. In another paper on the same topic, Tamminga & Hindriks (2019) argue that we can distinguish between logical reduction and ontological reduction. The latter implies that collectives are just sets of individuals, while the former implies that statements about collectives are logically equivalent to statements or conjunctions of statements about individuals. Tamminga & Hindriks (2019) argue for what they call 'logical irreducibility' of RGIs. Hence, I will assume throughout the paper that Hindriks (2019) argues as well for logical reducibility.

Hindriks (2019) provides two arguments for which we should understand the DJF as an irreducible obligation: firstly, the group bears an irreducible collective obligation because only the group can avoid the harm from occurring. Secondly, the individual lateral obligation of stage one has an irreducible collective content.

Concerning the first argument, first of all, consider the scenario by Feinberg (1968):

People lolling on the beach: a thousand people are lolling on a beach while one swimmer is drowning and shouts for help in a voice audible to the entire group of the lolling-people. Nobody intervenes and the swimmer is left to drown.

In this example, the responsibility to save the drowning swimmer is a reducible collective obligation: every individual able to save the drowning swimmer bears an obligation to do so (Feinberg 1968, 683-684; Hindriks 2019, 210). In cases such as offshore wind, things are different. The single individual cannot be considered as having an individual obligation to save the children, otherwise, we would have to say the same

for Solitary helplessness:

Solitary helplessness: we are in offshore wind. There is only one adult on the beach when the wind turns, and she can save only one child. (Björnsson 2020)

In solitary helplessness, we do not say that the adult has a moral obligation to save the entire group of children nor to contribute to save the group of children. The adult here is not able anyhow to prevent the collective harm or to bring about a relevant contribution to prevent it.¹² Thus, in offshore wind, the persistence of the obligation to save the group of children depends upon the presence of the other agents.

Regarding examples such as offshore wind, we have two options: (1) the individuals are morally obliged to bring about their relevant contribution to save the children and (2) the RGI as such is morally responsible to prevent the children's death. Hindriks (p. 210) argues that we should opt for (2). The point is that individuals have no motivation to bring about their specific contribution of the collective efforts if the others do not – just as in solitary helplessness the agent has no motivation to bring about a contribution to the collective effort if other people are not there to collaborate. The very existence of the obligation is based upon a joint reason that exists only insofar as there are enough individuals that share the same joint reason. Therefore, since the obligation is dependent upon the presence of the group, it is not reducible to individual moral obligations.

To support this latter claim, Hindriks (p. 210-212) offers a second argument. If we examine the content of the individual obligations to group up in phase one, we can see that the content of these individual obligations is irreducibly collective. Firstly, the obligation to ask others to contribute is dependent upon the others having the same obligation. Insofar as there are other agents with the same individual obligations, the individual obligations come into existence. Secondly, the obligation to answer affirmatively is conditional upon the others' willingness to help. One is allowed to refuse to help when the group is not likely to succeed (e.g., three persons would be needed but only two are available). Therefore, Hindriks argues that the collective obligation has to be logically prior to the individual ones: the content of the individual obligations to join forces derive from an irreducible collective obligation.

In this subsection, we saw why, for Hindriks, the DJF is an irreducible collective obligation. Firstly, because the reason to act is joint and it persists only insofar as the collective is present. Secondly, because the individual obligations to group up derive from the collective one that is logically prior to them. I will now conclude the presentation of Hindriks' argument in the next sub-section explaining his rejection of AT and the refusal of the so-called Irreducibility thesis.

1.2 Agency thesis and Irreducibility thesis

The main problem to ascribe irreducible responsibilities to RGIs is AT. To state the problem again, according to AT, only moral agents can bear responsibilities: however, RGIs do not seem to meet the agency requirements (Hindriks 2019; Björnsson 2020; Smiley 2017, 24-25). Therefore, to argue that RGIs can bear moral responsibilities it is either necessary to refuse AT or to argue that RGIs are moral agents. According to Hindriks, cases such as offshore wind question the validity of AT. Since such cases question the so-called 'causal condition' and the 'epistemic condition' forcing us to revise them, we should revise AT accordingly allowing for RGIs to bear irreducible collective responsibilities.

¹² This is justified by 'the causal condition of moral responsibility.' I will come back to this point later.

To consider an agent responsible, we must consider her as able to prevent the harm being brought about. This is the so-called ‘causal condition.’ However, as mentioned before, in offshore wind, no single agent can prevent the outcome from occurring but only the group by suitably combining their efforts. Hindriks proposes to reconsider the causal condition because, as offshore wind shows, there are cases in which multiple agents can prevent an outcome only by suitably combining their efforts (revised causal condition). Moreover, to conceive an agent as morally required to act, she should be conscious of the harm that is happening. In cases such as offshore wind, however, it is not enough that an agent is aware of the harm if the others are not. They must be mutually aware of the impending harm (revised epistemic condition). According to Hindriks, the revision of these two conditions requires the revision of AT. Cases similar to offshore wind require us to revise the causal and the epistemic condition on a collective level. Hence, we should refuse AT and allow for collections of agents to also bear moral responsibilities:

AT revised: the only entities that can bear moral responsibilities are moral agents and collections thereof (Hindriks 2019, 212. *Emphasis added*).

Consequently, if we refuse AT in favor of a revised AT, we can simultaneously argue that 1) RGIs are non-agential groups, and 2) they can bear irreducible collective obligations.

Nonetheless, Hindriks (p. 211-214) specifies that there is an important difference between obligations of agential and non-agential groups. In the RGIs, the collective obligation is grounded in the persistence of the individual ones. More specifically, for RGIs, the irreducibility thesis (hereafter, IT) is false.

Irreducibility Thesis (IT): A collective can be responsible for an outcome without any of its members bearing a correlative responsibility¹³ (Hindriks, p. 211).

Hindriks here argues for a central point: even if the collective obligation is not reducible to the individual ones, the collective obligation subsists only insofar as the individual obligations do. To see this, consider even stronger offshore wind:

Even stronger offshore wind: we are in offshore wind. The individuals passed stage one of the DJF and they are carrying the boat to the sea. While they are doing that, the wind suddenly becomes stronger and makes the two adults fall. Adult A breaks one of her legs and adult B breaks one of her arms.

Hindriks argues that here the collective obligation would be defeated by the individual obligations being given up. If two of the three adults of the group are needed to help the drowning children and two of them cannot help to bring about the action, then, the collective responsibility is defeated. Thus, IT seems to no longer apply with respect to RGIs.¹⁴

¹³ For correlative responsibility here we understand a responsibility that is dependent upon the existence of another one.

¹⁴ Here, it seems that Hindriks (2019) is drawing a similar point that Hindriks and Tamminga (2019) draw. It seems to me that Hindriks (2019) here is drawing a similar point. The collective responsibility is dependent upon the ontological existence of the individual ones. If the individual ones disappear (i.e., if there are less adults or one breaks his leg, etc.), then the collective responsibility also disappears. The kind of irreducibility for which Hindriks (2019) seems to argue throughout his paper is instead logical irreducibility: the statements we ascribe to the group are necessarily collective, thus making the collective obligation irreducible to individual ones.

This is an important clarification because in agential groups things differ. List and Pettit (2011) argue that excuses that apply to individual members of a group do not necessarily apply to the group itself. For a recent example, consider the role of the UK government during the Covid-19 pandemic. Even if for a considerable amount of time the UK prime minister was infected by the virus and he could not respond to his duties, the government was still responsible for the situation in the country. In RGIs, as even stronger offshore wind shows, IT is not valid: the irreducible collective obligation “exists only to the extent its members are in a suitable position to discharge it” (Hindriks 2019, p. 218).

To summarize this section, I showed how Hindriks relates his argument to two important theses concerning group responsibilities, AT and IT. On the one hand, he revises AT because it is not adequate to account for cases such as offshore wind. On the other hand, he refuses IT because the collective obligation of non-agential groups, while being not reducible to them, subsist only insofar as the individual ones do.

2 The Duty to Join Forces and Accountability

In this section, I will move to present a first challenge to Hindriks’ argument. Hindriks argues to have established a group obligation that is irreducible to the individual members. However, he still rejects the irreducibility thesis because the members carry a correlative responsibility to join forces. I will now argue that Hindriks’s argument is not compatible with a second kind of irreducibility either: irreducible accountability. I argue that, even assuming the validity of Hindriks’ (2019) argument, accountability in the DJF fully relies on the individual level. To show why this is the case, I evaluate three possible scenarios for which the group fails to comply with its obligation. I aim to show that we might have reasons to consider one or more individuals accountable, but never the group.

a) The group has no causal power to prevent the harm. In a scenario where the group has not the appropriate capacity to avoid the occurrence of the harm, it is unlikely that we hold the group responsible. For example, consider CPR:

CPR: we are in offshore wind. The difference is that, because of the wind, some high waves formed. The group of adults notices that some of the children are already drowning. They coordinate their action and take all of the children on the boat. However, three of them were already not breathing. They could have been saved if one of the adults had done a first aid course, but none of them had.

Assuming that taking a first aid course is not a moral obligation, we cannot hold the group responsible for the children’s death. Of course, if at least one of them thought about the benefits of such a course, the three children could have been saved. But this does not mean that we would blame or hold the group of adults morally responsible for the three children’s death. Therefore, the lack of a relevant ability to prevent the harm does not seem to be a justification to hold the group accountable for what could happen in the second stage.

b) The group obligation is defeated. As I mentioned before, Hindriks argues that the collective obligation can be defeated by individual ones being given up. As we have seen in even stronger offshore wind, in case of something happening out of the RGI’s control, the group obligation would be defeated. Interestingly, in case the collective obligation is defeated for something for which the individual is blameworthy, we would still not hold the collective responsible but only the individual as second thoughts shows:

Second thoughts: we are in offshore wind. The boat is bigger, and we need the entire group of adults to pick up the boat and save all of the children. After a successful first stage of the DJF, while the three adults are taking the boat, adult c has second thoughts and runs to save her child.

According to Hindriks, we have a collective irreducible responsibility, and yet, also in this case we would only hold adult c blameworthy for preventing the joint action that would successfully save the group of children. Postulating an irreducible group obligation does not seem to group accountability either here.

c) The collective fails in saving the drowning children. Finally, consider unlucky event:

Unlucky event: we are in offshore wind. After a successful stage one of the DJF, two of the three adults are running with the lifeboat on their shoulders to the sea. However, while doing that adult B slips and the lifeboat is falling. Adult C tried to jump to catch it before it fell, but she could not reach it. Adult A could not sustain the weight of the boat alone and the boat was destroyed. The adults run to save their own children.

In this scenario, the group fails as a group to comply with its obligation. We could therefore hold the group responsible for the harm. However, it is difficult to imagine that this would be the case. In moral philosophy, an important claim is that the will, the reasons, and the conditions under which the agent(s) perform the action are important to ascribe their responsibility (Strawson 1962; Scanlon 2008). In stage one, we already established the good will of the agents. They want to help the children and prevent them from dying. Postulating a group obligation of a group of agents of which we already established their goodwill and willingness to prevent the outcome does not seem to lead to group accountability.

Thus, if the group lacks a relevant ability, if the individuals would not join the collective efforts, or if the group fails while trying to save the children, the group's action seems to be morally irrelevant. Quite differently, imagine a corporation that committed itself to reduce its environmental impact. If the corporation is not able to reach what was stated, it would be likely that the group will be considered accountable for the outcome: (a) it is likely that the group has the relevant abilities to reduce their emissions; (b) it would not be an excuse that one of their members was unable to contribute; (c) the corporation's good intentions would often not be enough. Our attitude changes drastically only for one reason: we tend to consider the corporation as an autonomous agent, and, therefore, we expect it to act on certain considerations of good and bad as a collective.

However, as shown above, things change drastically with RGIs. I argued that in the three scenarios above, we would not hold an RGI responsible because the group's action does not seem to be relevant from a moral point of view. In RGIs what seems to be morally relevant instead lies on the individual level. Individuals are accountable for their decision to accept or not to collaborate to the joint action.

3. Irreducible Responsibilities Without AT

In the previous section, I have argued that Hindriks (2019) fails in postulating irreducible accountability. In this section, I move to a second criticism. I argue that without AT, Hindriks' argument seems to fail in postulating an irreducible obligation.

As we have seen, a crucial condition for an agent a to be responsible for a harm x is that she must be able somehow to prevent it. In offshore wind, there are two other individuals. Agent a is obliged to join forces with them thus preventing the harm. However, on the one hand, consider that the presence of the two individuals might not be enough. For example, if the other two adults are physically impaired, the group

would not be able to prevent the outcome. On the other hand, we should also acknowledge that the presence of two more people is not necessary for agent a to be morally obliged to save the children. She might be individually able to save the children. Thus, what is necessary for agent a, to be morally obliged to take action, is that she has an acceptable way to save the children's life.¹⁵

In offshore wind agent a is morally obliged to take action. She must join forces with the other adults and then maybe row on the lifeboat while another adult does the same. According to Hindriks, this obligation is irreducibly collective. However, consider alternatives:

Alternatives: we are in offshore wind, but there is also a motorboat that only adult a knows how to use. Adult b asks adult a to help her save the children's life. Adult a quickly refuses the offer and uses the motorboat to save the children's life.

According to Hindriks' argument, adult a should be taken to having two distinct moral obligations. On the one hand, she is subject to the DJF of the group of adults. Adult a is morally obliged to agree to help when adult b asks for it. On the other hand, adult a would have an individual moral obligation to save the children's life through the motorboat. Therefore, it seems that (a) in the scenario, adult a is subject to two different moral obligations: an individual and an irreducibly collective one; (b) taking Hindriks' argument strictly, adult a would commit wrongdoing by refusing to help.

About this latter point, I think it would be an unfair criticism of Hindriks' argument. He did not account for examples like alternatives and surely there is room to clarify how the DJF would deal with such a situation. Therefore, we can exclude that adult a commits wrongdoing using the motorboat and refusing to join the joint effort as long as she can save the children. What might be instead worth looking at is (a). For Hindriks' argument, adult a must be subject to two different obligations. If agent a would be subject to two equal statements, the DJF would be reducible to an individual obligation, thus, problematizing Hindriks' argument.

The two obligations in alternatives seem to be equal for at least two factors. Firstly, complying with one of the two obligations would make the other disappear. Saving the children through the motorboat would remove the RGI's obligation to save the children. Secondly, it is not morally relevant whether adult a complies to the obligation as a group member or through the motorboat. The two obligations are equally strong since they refer to preventing the same harm through equally acceptable means. Thus, agent a can freely choose what she thinks is better to save the children's life.

The only difference between the two obligations is that, in one of the two, there is the contribution of other agents. But is this enough to claim that agent a is subject to two unequal statements? Through his argument, Hindriks argues that it is the case. The mere presence of other agents makes the obligation an irreducible collective one. In alternatives, there must be something in the presence of other agents that alters the obligation that is instead missing in the contribution of the motorboat. The causal contribution brought by agent b and by the motorboat can be considered equal to agent a. Hence, the difference must lie somewhere else. Surely, one important difference between the presence of agent b and the motorboat is the fact that agent b is indeed an agent. Thus, this might be the factor that generates an irreducible collective obligation.

15 I do not go into considerations of what counts as an 'acceptable' way to save the children's life. Consider however that she would probably be excused to not save the children if she would have to risk her own life. Considerations alike are possible. See (Hindriks 2019; Björnsson 2020).

In the first stage of the DJF, the fact that agent b is an agent might be important. The two individuals have to agree to help before taking action. Instead, in the case of the motorboat, the agent can autonomously decide whether to save the children or not. However, this difference does not matter for the subsistence of the obligation. Hindriks himself argues that the obligation to save the children – that according to him is collective – is prior to the lateral responsibility to join forces. Hindriks' argument is built upon the idea that the mere presence of the individuals generates an irreducible collective obligation. It does not matter, for the persistence of the obligation, whether the agents are going to agree or not, whether agent a is a master hypnotist and might ask an unwilling agent b to help, etc. Reading through Hindriks' argument, the existence of the collective obligation is only based upon the relevant contribution that agent b might bring, and which will result in essentially combining with agent a's contribution to then prevent the harm. But we have assessed that the motorboat would bring a similar causal contribution and that the difference must lie somewhere else. However, the fact that agent b is an agent does not seem to be relevant for the obligation to occur, and it remains unclear where the difference would lie.

Before concluding, it might be worth noting that a possible way to argue that the two obligations are different would be to claim that in the first stage of the DJF, agent a and agent b would constitute themselves as an autonomous collective agent. We should assume that they would institute for themselves a structure and a decision process and that the group would then meet the agency requirements. If we assume that this is the case, we have reasons to believe that agent a is subject to two unequal claims. On the one hand, she is subject to an individual moral obligation that fully weighs on her agency. On the other hand, she is the indirect subject of a moral obligation since she is a member of an agential group upon which the obligation weighs. Thus, in this case, it seems possible to argue that the two obligations would not be reducible to one another. But this possible path of argument would occur respecting AT, and not refusing it, as Hindriks' argument aims at instead.

Thus, it seems to me that a much more straightforward interpretation of alternatives is that adult a has two ways to prevent the harm and comply with her single obligation to save the children. But if agent a in alternatives is subject only to one obligation, it means that the DJF is reducible to an individual obligation, thus questioning the validity of Hindriks' argument.

Conclusions

In this paper, I presented Hindriks (2019). Hindriks argues that it is possible to have irreducible group responsibility for random groups of individuals. I argued for two points. Firstly, I argued that while the DJF leads to postulate an irreducible group obligation, it would not lead to irreducible accountability. Secondly, I argued that Hindriks' argument at its roots would be problematic. I argued that it is unclear why the mere presence of other agents should alter the obligation.

To make his argument appealing, I believe that Hindriks should answer both my critical points. Firstly, he should make clear the reason for which joining forces with other agents would necessarily make the obligation irreducibly collective, if what is needed for the obligation to occur is only their relevant contribution but not their status as agents. Secondly, accounting to this criticism would still leave open another issue. Hindriks seems to postulate an irreducible group obligation for which the group is not accountable. It is unclear what Hindriks' argument would add if accountability would still lie on the individual level. I believe that accounting for these two challenges can be complex. Trying to ascribe responsibilities without establishing agency might be an unrealistic task.

References

Björnsson, G. (2014). Essentially Shared Obligations. *Midwest Studies In Philosophy*, 38(1), 103-120.

Björnsson, G. (2020). Collective Responsibility and Collective Obligations Without Collective Moral Agents. *The Routledge Handbook of Collective Responsibility*, 127-141.

Held, V. (1970). Can a Random Collection of Individuals be Morally Responsible? *The Journal of Philosophy*, 67(14), 471.

Hindriks, F. (2019). The Duty to Join Forces: When Individuals Lack Control. *The Monist*, 102(2), 204-220.

List, C. & Pettit, P. (2011). *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford University Press.

Scanlon, T. (2008). *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, Mass. Harvard University Press.

Singer P. (1972). Famine, Affluence and Morality. *Philosophy and Public Affairs* 1(2), 229-43.

Smiley, M. (2017) "Collective Responsibility", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.).

Strawson, P (1962). Freedom and Resentment. In *Proceedings of the British Academy*, Volume 48:1962. pp. 1-25.

Tamminga, A., Hindriks, F. The irreducibility of collective obligations. *Philos Stud* 177, 1085–1109 (2020). <https://doi.org/10.1007/s11098-018-01236-2>

Williams, B. (1981). Persons, character and morality. In *Moral Luck: Philosophical Papers 1973-1980* (pp. 1-19). Cambridge: Cambridge University Press.

Wringe, B. (2010). Global Obligations and The Agency Objection. *Ratio*, 23(2), 217-231.

Promoting freedom in religious education

Rodrigo Bustamante

1 Introduction

1.1 A policy dilemma

The Cornelius Haga Lyceum (CHL) in Amsterdam is the only Muslim secondary school in the city. It has been the focus of controversy for some years. On one hand, the school has been connected with extremist groups: a school board member—who has since resigned—showed support to extremist groups on social media. A second element of the controversy regards the school's religious practices: boys and girls are required to attend class separately, girls are required to cover all parts of their bodies except hands and face, they may be called to mandatory prayer, and are expected to fast during Ramadan¹⁶. Popular media has shown concern that children could be “indoctrinated” through these religious practices, close themselves off from society, incorporate harmful gender roles, and be exposed to intolerance toward homosexuality. In response, the government has opened investigations and some politicians have proposed increased regulations. In turn, this has made other religious schools of various affiliations weary that increased regulations would reduce their capacity to practice religion at school¹⁷.

For the purpose of this article, I will leave the first element of the controversy aside. I will approach only the second, which invites significant questions regarding the liberties¹⁸ that the state protects and promotes. The situation seems to present a policy dilemma where two types of freedoms are in tension. First, the children's freedom of religion: the freedom to adhere to one's religious mandates. Second, the children's freedom from interference: the freedom to choose any version of the good life that one pleases without being indoctrinated into a particular conception of the good. The policy dilemma can be described as follows. If the state is concerned with children's freedom from interference, one would argue it should stop funding CHL or close it altogether. But if the state is concerned with children's freedom of religion, one would argue that the school should remain open and perhaps also funded.

A liberal state has an interest in safeguarding both of the freedoms above. More generally, a liberal democracy is ruled by a principle of liberal neutrality (PLN). This principle requires the state to remain neutral to conflicting conceptions of the good (Snik & De Jong, 2002). According to the PLN, the state should not promote any particular version of the good—for instance, a religious one—it should limit itself to protecting and facilitating freedoms. For a state that is ruled by the PLN the CHL policy dilemma seems to involve two types of freedoms that are in tension.

1.2 Argument outline

My goal is to present a framework by which the tension in the above policy dilemma can be resolved. I argue that the tension can be dispelled by adopting a conceptualization of freedom that is based on Christman (1991). This framework focuses not on what children are free to do or free from but rather on how they develop their preference to act in a certain manner. More specifically, in order to act freely, the process by which one develops a preference to act should respect one's autonomy.

16 <https://www.rtlnieuws.nl/nieuws/nederland/artikel/4634376/minister-en-aivd-slaan-alarm-over-islamitische-middelbare-school>

17 <https://www.nrc.nl/nieuws/2020/01/21/vonnis-haga-lyceum-test-voor-onderwijsinspectie-a3987671>

18 Throughout the essay I use the words liberty and freedom interchangeably.

The article is structured as follows: In section 2 I make further distinctions between the two types of freedom that appear in the policy dilemma. First, the freedom to practice one's religion according to its precepts and doctrines—children are free to grow into their religious community, even if their religious practices are somewhat controversial from the outside. Second, the freedom from interference in choosing one's own version of the good life—children are free to grow as individuals without religious authorities dictating what they should value.

In section 3 I present two different frameworks to conceptualize the policy dilemma. The first is due to Berlin (1958) and distinguishes between positive and negative freedoms. The goal of presenting this framework is to draw a parallel between the policy dilemma on the hand, and Berlin's paradox between positive and negative freedoms. This framework is appropriate to characterize the tension and understand it, but not to solve it. In contrast to Berlin's, I present a second framework due to Christman (1991). This alternative, compared with Berlin's, dispels the paradox by identifying freedom with autonomy.

In section 4 and 5 I present possible objections to Christman's framework. The first one is that, since children are not fully autonomous the framework cannot require that their autonomy is respected. My response is that, if one considers that individuals begin to develop their autonomy in childhood, the conditions stated by Christman's framework still apply. A second possible objection is that it may not satisfy the real demands of CHL or some of the religious schools. These could object that the requirement to respect autonomy is too demanding—it means giving up on precisely the kind of practices they want to impose. My response to this second objection is that, for such cases, it may not be possible to solve the tension in the policy dilemma. If a school does not want to concede to basic principles of liberal society more broadly—if it does not accept that the PLN holds for society as a whole—a policymaker may not be able to solve the dilemma. In such a case she could not satisfy both the schools demands regarding children's freedom of religion, and the demands of children's freedom from interference.

2 Two freedoms in education

In this section I further describe the two freedoms that seem in tension in the CHL policy dilemma. My goal here is to make clear what freedoms are involved, and how they are indeed different. I first present freedom of religion and describe what this freedom requires from the state. I then go on to do the same for freedom from interference.

The CHL case seems to place two types of freedom in tension. But are these two types in fact different? One could object that freedom of religion is actually a particular instance of freedom from interference. After all, we can say that an individual would be free to practice her own religion if she is free from the interference of outside pressures—such as a state that pushes for a particular religious affiliation. For instance, Muslim followers in CHL are not forced into Christianity through interference from the Dutch the government. But is this enough to claim that they are free to practice their religion? In fact, an individual may in fact not be forced to abandon his religious affiliation, and still be unable to practice it. Consider for instance, that CHL were to be closed, and that no Muslim school were to be present in the city. There would then be no possibility for a religious education for the Muslim youth. Exercising one's freedom to be educated in the teachings of the affiliation would likely be much more difficult. The freedom to effectively practice a religion according to its teachings, therefore, seems distinct from the freedom to be interfered with in one's beliefs. Another objection to separating the two types of freedoms may be the following. One could agree that these two freedoms are of importance for children, but object that freedom of religion can be practiced outside of school, and that there is no need for this freedom to be expressed in this specific environment. Following this objection, the liberal state could focus on promoting freedom from interference in a particular sphere,

and freedom to practice religion in a different one. However, individuals do not exercise freedom of religion in a void, they do so within a cultural group—a community that holds one's same values and practices. According to Kymlicka (1995), cultural membership is needed for the practice of individual freedoms. Even more, some versions of this freedom require these values to be present in all spheres of lives. Hewer (2001) for instance argues that, within an Islamic worldview, there are no secular subjects, and all aspects of life are permeated by religious values. This makes education an essential sphere in which freedom of religion is practiced.

Now, what does the children's freedom of religion require from the state? Does it imply that the state should tolerate faith schools, or that it should promote them? In other words, does this freedom require the state to allow them to remain open, or to fund them publicly? This may depend on how the PLN is specified. If the PLN is stated as a requirement for the state to simply protect individual liberties, then freedom of religion would seem to require only the toleration of faith schools. If in contrast the PLN is stated as a requirement for the state to promote individual freedoms, then one could argue that it ought to fund religious schools as well. Throughout this article, I am not making a determinate choice between these two formulations of the PLN. The reader could hold one or the other. It is important to notice, however, that the choice of formulation of the PLN does not change the premise I try to establish here: freedom of religion is a distinct freedom involved in faith schooling—whether this means a requirement of toleration or promotion from the state.

My claim here is that, minimally, a state ruled by a PLN is required to tolerate faith schooling. One reason for this becomes clear by making a distinction between primary goods and secondary goods (Rawls, 1993; Snik & De Jong, 2002). Education is usually seen as a primary value, one that is widely accepted by different members of society, no matter their cultural background or the other values that they hold. This is, firstly, because it enables children to participate in society as future adults. Education does not need to be understood as formal schooling, but as any form of relatively structured rules and norms that guide the upbringing of children into society. Indeed, the importance parents give to education seems generally independent from political, cultural or other values that they hold. Parents from different groups in society could disagree on many of their values, but likely agree on the fact that their children's education has fundamental importance. It is therefore difficult to think that any given conception of the good would negate the importance of education. In contrast, specific forms of education, such as faith schooling, can arguably be regarded as secondary values. Secondary values may be dependent on different cultural and political traditions. As Snik & De Jong (2002) exemplify, charity and empathy may be prominent in a religious conception of the good, while a secular tradition may put higher value on personal need-satisfaction. Since secondary goods are specific to particular social groups, they may conflict with each other. We can see this in the context of the CHL: secular outsiders may put higher value in freedom from interference, while religious insiders may place higher value in freedom of religion. This may be the case even while both agree on the primary value of education in general. If we take it as a premise that the state is ruled by a PLN, the requirement of neutrality implies that it should not promote any particular secondary value. If it did, it would be favoring a particular group. At the same time, the neutrality of the PLN also implies that the state should tolerate opposing secondary values in society. On one hand, some individuals may hold secular and atheist secondary values (perhaps more in line with freedom of interference), while others would hold religious secondary values. Under the PLN, the state is required to tolerate these two conflicting secondary values.

In the previous paragraphs I attempted to convince the reader that (1) freedom of religion is a distinct concept from freedom from interference; (2) freedom of religion is relevant in the context of CHL and in religious schooling in general; (3) the state is required to either promote or tolerate expressions of this freedom; and (4) toleration is a minimal requirement, even when it may lead to tensions between groups holding divergent values. I am consciously leaving the fourth statement largely undefined; precisely how much conflict should or can be tolerated is a significant question, but I am putting it aside for the purpose of this article. My point is merely that some degree of conflict is implied by toleration of diverging secondary values.

Let me now move on to freedom from interference. Liberal neutrality is usually invoked against interference from the state advancing any particular conception of the good (Snik & De Jong, 2002). But the state is not the only possible source of interference. Other groups or individuals may forcefully push to advance their particular secondary values. These include institutions like schools or religious associations and churches. Freedom from interference is not restricted based on the source of the restriction. The interference may not necessarily come from distant groups or individuals, it may also come from members of one's own family or cultural group. The requirement of freedom from interference applies to these cases as well. Under this freedom, children ought to be protected from interference even from their own schools and families.

Here a qualification is needed. One may say that freedom from interference cannot be an absolute requirement when referring to children. Parents certainly should be allowed to interfere in some way in the upbringing of their children. I discuss this concern in more detail in section 5. For the moment, the reader may accept at least the less-controversial requirement of no interference from institutions like the church or school.

3 Two frameworks to understand the freedoms involved

In the previous section I attempted to establish two premises:

(1) The state is ruled by the PLN.

(2) Freedom of religion and freedom from interference are distinct freedoms, and they are in tension in the CHL case, and in religious schooling more generally.

We can now address the main question: what considerations should a policymaker have with regard to cases like CHL if they aim to safeguard both of the freedoms that seem at tension?

Liberty may be understood from two different perspectives (Carter, 2003). On one hand, we may focus on the content of liberty, or what an individual is free to do. From this perspective, we may think that the children in CHL have the freedom to receive a religious education. A different perspective focuses instead on the process by which liberty obtains, or how an individual develops the will to act in a certain way. Under this perspective, the children in CHL are free to receive a religious education, as long as this respects a particular procedure.

In the following two sections I present these two frameworks in more detail. The first, content-focused framework due to Berlin (1958) is presented in section 3.1. By presenting this framework I attempt to draw a comparison between the policy dilemma between freedom of religion and freedom from interference, and Berlin's concepts of positive and negative freedoms. This parallel is useful in understanding how these freedoms are at tension, but it is still insufficient to resolve the dilemma. In section 3.2 I present a content-neutral or procedural framework. This framework describes freedom as requiring autonomy and allows to present the dilemma in a new light. In doing this, the tension between freedom of religion and freedom from interference can be resolved.

3.1 Content-focused view: positive and negative freedom

According to Berlin (1958) negative liberty defines the area in which the subject should be left to do whatever she is able to do without interference from other persons. Positive liberty, instead, defines the source of control that determines a person to act in a certain manner (Berlin, 1958). The distinction between these two forms of liberty has since been widely incorporated into political philosophy (Carter, 2003).

Being negatively free involves an absence of obstacles to realize what one could realize. These obstacles, as Berlin points out, are external. Being incapable of performing a certain action or becoming a certain person due to internal factors does not make one negatively unfree. The inability to do or become comes from other persons deliberately interfering with one's ability. A liberal state accepts that children attending religious school have the freedom not to be interfered with in choosing any conception of the good. Consider a person first exposed to her sexuality in her teens. Suppose she realizes that she is attracted to the same sex, and that she decides to pursue this attraction. We say she is negatively free if there are no interferences between her and the realization of this desire. We may say she is negatively unfree if at school she is exposed to intolerance and taught this to be a "mistaken" conception of what is good. This example makes clear a parallel between freedom from interference and negative freedom. With regard to religious education, we may say that children are negatively free if they are not interfered with in determining their choice of affiliation or if they become religious at all. Freedom from interference in the CHL policy dilemma can therefore be described as a negative liberty in Berlin's framework.

Positive freedom, on the other hand, involves the capacity to act in such a way as to fully take control of one's life through self-determination. It involves realizing one's most fundamental and true purposes. Here, liberty is concerned not with the area of control, but with its source. "For the musician, after (she) has assimilated the pattern of the composer's score and has made the composer's ends (her) own, the playing of the music is not obedience to external laws, a compulsion and a barrier to liberty, but a free, unimpeded exercise" (Berlin 1958, 14). One can take this example back to religious freedom. For many children receiving a religious education is essential to developing a relation with their community. They may want to realize themselves by living a devout life and eventually joining their community in all their religious practices. If this is the case, their impulse to follow praying times, follow requirements of clothing, and stop themselves from acting on their homosexual desires, simply because their truest wish is to live by the religious dictum. Note that the freedom of religion that is involved in the policy dilemma, therefore, cannot be described as negative freedom. An individual may find no impediment to live religiously and, in that sense, be negatively free to live a religious life. In contrast, being positively free requires that her wishes are truly her own, and that she effectively realizes them.

So far, I have described the two freedoms involved in the policy dilemma as positive freedom (religious freedom), and negative freedom (freedom from interference). In his article, Berlin (1958) pointed to a fundamental problem between these two types of freedom. This takes the form of a "paradox", in Berlin's words, by which positive freedom carries with it a danger of authoritarianism, and therefore a direct threat of interference over one's negative liberties. According to Berlin, the paradox comes about in two steps (Carter, 2003). First, theorists of positive freedom assume a divided self. One that is truest to one's ultimate desires, aligned with self-realization, and commonly thought to be the rational or higher self. In contrast, the lower or impulsive self is dominated by elusive passions. To illuminate this point, consider a smoker who is genuinely wishing to quit the habit. As she finishes her beer and looks out to the garden, she is overwhelmed with the desire to light a cigarette. At the same time, she has not smoked in weeks, and she fiercely wishes

she could improve her health prospects and stop smoking altogether. If these conflicting desires can be thought of as belonging to different selves, one seems to be truer to her desires, since she wishes she would not even want to smoke. In a second step toward the paradox, this separation of selves can be extended to different individuals. Under this premise, some may be in a position to know what people's rationality or truest realization consists of. In the smoker example, think for instance of a policymaker who claims to know that most smokers, even while they find it difficult to quit the habit, would actually prefer to do so because it aligns with their truest preferences. In the CHL case, think of what proponents of closing the school altogether may claim: that even if children would manifest wanting to receive religious education above all, they don't really know what they want because they are not responding to their truest self. This puts the proponent of closing the school in a position by which she claims to know more about the children's true preferences than the children themselves. According to Berlin, this presents a paradox that can justify some individuals interfering with other's actions, allegedly in the name of their liberty.

What does this mean for the CHL case? What does it say about the mandate of the liberal state? Berlin's distinction between positive and negative liberties allows us to frame the tension we observed between freedom of religion and freedom from interference in a new light. We can therefore describe the tension between the two liberties and the threat that the dilemma presents. In an attempt to promote children's positive freedoms, the liberal state could be justifying some individuals in interfering with their negative freedoms.

Although Berlin's paradox of positive and negative freedom helps describe the policy dilemma, it does not give tools to solve it. In the next section, I contrast Berlin's content-focused conception of freedom with the content-neutral conception of freedom due to Christman. This second framework will help us move from describing the policy dilemma, to providing a possible solution to it.

3.2 Content-neutral view: freedom as procedural autonomy

Christman (1991) attempts to resolve the tension generated by Berlin's paradox by developing a different concept of positive freedom. Here, positive freedom is not defined by the content of an action, but by the process that brings it about. In other words, a particular action is not freedom-preserving or freedom-diminishing because of its content (CHL enforcing a dress code, fasting, or exposing children to intolerance toward other social groups). Rather, an action is free or unfree depending on the process by which the agent comes to it.

To understand the distinction between these approaches, Christman presents an example. Consider a woman who acts according to her fundamentalist religious beliefs, one of which dictates that a wife ought to be subservient to her husband and should have only a domestic role. Suppose she developed her values within an oppressive environment: her father was an authoritarian figure. The man she is married to now is abusive. Suppose that, today, her truest desire is to be a devout religious follower and remain subservient to her husband despite the abuse. Suppose further that nothing is stopping her from escaping the situation: her social relations would not be severed, and her family would materially support her if she made the decision to leave. Her desire to remain in the situation she is in is driven by her desire to be a proper follower of the religion.

Berlin's content-focused view of her freedoms tells us that she is seeking to be positively free by living a devout religious life. At the same time, she seems to be negatively free: no one is stopping her from leaving her home and starting a new life. Granted, the husband would not be delighted, but she could still do it

without impediment. Despite this, are we intuitively willing to accept that she is indeed free? She may be truly free in her choice to stay, but she may also be conditioned by her past experience. Christman would reply that she is indeed not free, and that to understand why, we need to observe the process that brings about her preference.

According to Christman's account, for her to be acting freely, her character needs to have been constituted in a particular manner. What we need is an account of how her preferences are formed. Christman's strategy is to identify positive freedom with autonomy in this process. Generally, autonomy is a concept closely related to individual freedom, but it is distinct from it. Freedom concerns the ability to act to make one's desires effective, autonomy refers to the independence and authenticity of one's desires (Christman, 2018). In Christman's account, if an action is free, the agent's preference needs to develop through a process that respects her autonomy. Christman states the conditions for autonomy as follows:

1. The agent is in a position to reflect upon her choices and upon the changes to her preferences.
2. The agent is in a position to resist her preference change if she wants to.
3. Reflection and resistance are not the product of further factors that inhibit self-reflection, such as drugs or alcohol¹⁹.
4. Self-reflection, and the desires that result, are minimally rational in an internalist sense: there are no manifest inconsistencies.

On one hand, this framework captures the idea that one is free when one is the ultimate source of one's action. On the other hand, it guards against the dangers of authoritarianism. It achieves this by providing procedural conditions that a free action ought to fulfill. We see this by noting that, if an agent is in the capacity of meeting conditions 1-4, her actions are coherent with her positive as well as negative freedom. This invites further questions regarding the example above: did the devout follower, who does not want to leave her abusive husband, come to desire her current situation autonomously? Has her upbringing enabled her to self-reflect so that she became aware of her desires as they are today? Was she somehow threatened in the process of making this decision—for instance either by social and cultural isolation or by financial insecurity? Questions like these become relevant in identifying to what degree she is free in her choice.

For the policy dilemma in the CHL case, and generally for religious schooling, this means that we have conditions which a liberal state should impose on the school, if it is to safeguard both of children's freedoms. The students' choice is to receive a religious education through which they incorporate a particular conception of the good. Are they—to a reasonable extent—in a position to reflect on the significance of this choice? Do they have the space to consider different life paths they could take in a liberal society? Have they had the opportunity to resist this choice? How does the school relate to students who do not go on to fulfill devoutly religious lives? Questions like these are naturally very difficult to answer objectively. Even more, there may be no binary answers to such questions. These conditions may be fulfilled only in degrees. One should notice, however, that they are not impossible to answer reasonably. A procedural account of freedom seems capable of providing at least a roadmap for the policymaker to ask these questions and formulate policy decisions in ways that attempt to respect children's autonomy as a safeguard of their freedoms.

There is, admittedly, yet another difficulty. Not only are the conditions for autonomy hard to answer, but—importantly—they refer to children. Children are typically seen as not yet having developed a fully free, rational and autonomous agency, and parents have certain rights and obligations with regards to making decisions in their place (Brake & Millum, 2012). The next section briefly discusses this difficulty.

¹⁹ This is not to say that these effects cannot form part of an autonomous process of preference formation. If the decision to take them was made autonomously, then the process will indeed respect autonomy.

4 Education and children's autonomy

If we accept that children are not yet fully autonomous, how can we require that CHL respects their autonomy? According to Christman (1991), it is indeed possible to evaluate whether a certain childhood favors autonomy. He provides a backward-looking criterion for this: "If I look back on my past and see that much of my character was formed by educational and parental practices that I would not want to have been molded by (...) then I am not autonomous relative to those aspects of my character" (Christman 1991, 349). There are three reasons why such a backward-looking criterion can be problematic. First, one can only identify the lack of autonomy after the fact. Therefore, it is difficult to see how this could be used as a policy criterion. Second, Christman's criterion seems highly personal. Different individuals may indeed look back at their childhood through various perspectives; similar practices may be regarded differently depending on a wide range of factors. Take any particular child attending the CHL, for example; she may think of gender-separated classrooms as harmless and in line with her familial upbringing. This may especially be the case if she, as an adult, reaffirmed her childhood values and those of her environment. For a different child in the CHL, looking back at her experience may be completely different. Consider for instance a child who grew up to become a fervent atheist and who has completely changed her social environment in line with less religious values. Because the practices of CHL did not align with the values she developed as an adult, she may regard them as oppressive. One same practice, then, may be seen as autonomy-preserving or as autonomy-diminishing depending on the agreement to one's own current personal values; such a criterion seems to go against the kind of value-neutrality that is required by the PLN. Third, it seems that Christman's backward-looking criterion could only be applied after the individual has gained autonomy. If someone has had an autonomy-diminishing childhood, and is still constrained in this respect (autonomy conditions 1 to 4 are not met), how could she realize this? How could she look back on her upbringing, and be able to judge it as autonomy-diminishing? For her, Christman's backward-looking criterion seems to fail to identify the problem.

How, then, can a policymaker begin to evaluate if a religious school respects the autonomy of children, when they are not yet autonomous? A possible reply to this objection may go as follows. Education is a space in which children develop their autonomy, even if they are not fully autonomous yet. It seems important that this development is not hampered, but promoted, so that they eventually become autonomous adults. In other words, childhood and teenage years are a domain of preference formation. Individuals at this stage begin to form preferences of their own, and they can, to some extent, begin to do so autonomously. It seems reasonable to assume that some children do begin to think critically about themselves, their beliefs, and those of their family and community. Consider a particular student in CHL: even when most other students may accept the practices imposed by the school uncritically, it would not be surprising to see that one of them forms preferences against, for example, the class separation by gender. The key distinction I wish to make here is that children, even if not yet fully developed, have the potential to act autonomously. Think of the CHL student once again, and how she may attempt to reject some of the school practices. She may voice an opinion against mandatory praying hours, she may begin to prefer learning about other religions, or she may begin to challenge role models imposed by religious practices. How could the school react to this, and what influence would that have on her? If the school tries to mute her critical reflection, we would intuitively accept that her autonomy is being hampered. On the contrary, the school could give her the opportunity to find out more about other religions, while she reaffirms her commitment to her own. In this case, we would intuitively accept that her autonomy is being respected.

If we accept that, to some extent, children begin to show autonomy, the conditions for autonomy stated in section 3.2 may still apply. We only need to change their formulation slightly. Instead of asking “Have children formed their current preferences in accordance with the autonomy conditions?” we may ask “Could children develop their preferences in accordance with autonomy conditions?” If the autonomy conditions 1 to 4 are met, we may answer this question affirmatively. It seems, therefore, that to the extent that children begin to develop their autonomy and self-reflect critically, it is still possible to apply the autonomy conditions stated above.

5 Moderate vs fundamentalist challenges

One may raise a further objection to Christman’s framework. Consider that we accept the premises and a content-neutral approach to liberty. The liberal state may now call upon religious schools and demand measures to safeguard children’s autonomy. Practically, policymakers may instantiate this requirement in various ways. For instance, the school could make gender separation a requirement only in religion classes, plan field trips or sport competitions with other schools, monitor teaching material for intolerant content, make praying times optional, and openly teach about sexual orientation while discussing religious dictates. But—one may contend—doesn’t this miss the point altogether? Some of the schools, in fact, could not want to apply these measures, because they imply a defeat of their religious requirements entirely.

This is the point advanced by Burtonwood (2003), who claims that the challenge for liberals is not to respond to a “moderate” version of religious schooling, which already is in line with the basic principles of the liberal state. Instead, the true challenge comes from “fundamentalist” schools or communities, who may be unwilling to accept basic tenets of liberalism. According to Burtonwood, this puts liberals in a position in which they have to either bite the bullet and admit that liberalism is just one more conception of the good, being imposed by the state, or accept a radical form of pluralism and drop any conditions on religious schooling altogether.

The dichotomy stated by Burtonwood above seems mistaken. The response seems to confuse the role that liberalism, and the PLN in particular, has in moderating the challenges I am addressing in this article. Schools or groups who are entirely unwilling to accept any form of compromise that entails the respect of a PLN may indeed reject any policy that is based on the arguments drawn so far. But it does not follow from their rejection that liberalism is a conception of the good among others. Here, I am following Snik & De Jong (2005) in a similar response. The PLN does not define any particular conception of the good, it defines the scope that freedoms may take, and the limits required so that the freedoms of a particular individual do not reduce the freedoms of another. What the PLN advances and presupposes, is that individuals who have conflicting values and demand different freedoms, are willing to recognize each other as part of the same overarching social group, and to give each other’s liberties the furthest possible scope. In this sense, liberalism both enables and restricts freedom. It restricts the freedom of an individual when this may trample over another’s—regardless of their personal values. In doing so, it also enables the freedom of those who would otherwise be trampled on. The PLN, therefore, frames competing conceptions of the good, making their coexistence possible. It achieves this by conditioning their conflicting freedoms, for instance, on respect for individual autonomy.

Two points follow from the above description of the PLN: First, the fact that a religious school may refuse to accept the common limitations imposed by the PLN, is not a reason for other groups to abandon the PLN altogether, nor to believe that the PLN is simply one secondary value among others. Second, the school’s refusal does not mean that other members of society need to accept a radical form of pluralism in which no conditions are imposed on religious schools either.

If a policymaker is confronted with a situation like the one above, it may indeed not be possible to propose a policy solution that would be accepted by the most fervent religious school. If any particular group is unwilling to see others as members of the same society with an interest in setting mutual limits and enabling mutual freedoms, any dialogue is stalled. In such a case, it does not seem possible to find a principled solution to the tension.

6 Conclusion

Let me take stock. The case of the Cornelius Haga Lyceum in Amsterdam, and religious education more broadly, present an apparent policy dilemma. The children's freedom from interference appears to be in tension with their freedom to receive a religious education. A state that is guided by a Principle of Liberal Neutrality will need guidance on how to safeguard both of these freedoms.

To unravel the policy dilemma, I presented two frameworks. The first, a content-focused framework due to Berlin, allows us to describe the policy dilemma in terms of positive and negative freedoms. Their tension is described by Berlin as a paradox by which positive freedom presents a threat of authoritarianism to negative freedom. This framework is useful in describing the dilemma and its tension but is not enough to provide a solution. This is what the second framework I presented aimed to do. Christman's content-neutral approach identifies freedom with procedural autonomy. In this framework, the children are free not by merit of the content of their actions. Instead, they are free as long as the process by which they develop their preferences respects certain conditions of autonomy.

The exposition allows us to derive policy guidance. What the liberal state should require of religious schools such as CHL is that they protect and facilitate the development of children's autonomy. Practical examples of this may include limiting gender separation only to religion classes, increasing exposure to other schools, monitoring teaching material for intolerant content, and teaching about sexual orientation openly. In Christman's framework, autonomy requires, most importantly, that children are able to reflect upon changes in their preferences, and to resist these changes if they want to.

In the final two sections briefly discussed possible objections to Christman's framework. First, one may view children's autonomy as not yet fully developed. Children may not be in a condition to exercise the kind of self-reflection required by autonomy. I argued that children do indeed have the potential to act autonomously. The conditions stated by the proposed framework therefore continue to hold. The second objection regards the challenge of "fundamentalist" schools who may not want to adopt any requirement regarding autonomy. I argued that the PLN is not a particular conception of the good, but a framing rule to negotiate conflicting versions of the good. If any group—regardless of its values—is unwilling to partake in this mutual compromise and negotiation, there cannot be any principled solution to meet its demands.

References

- Berlin, I. 1958. "Two Concepts of Liberty." In Isaiah Berlin. 1969. *Four Essays on Liberty*. Oxford: Oxford University Press.
- Brake, E., & Millum, J. 2012. "Parenthood and Procreation". *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/entries/parenthood/#BeiPar>
- Burtonwood, N. 2003. "Social Cohesion, Autonomy and the Liberal Defence of Faith Schools". *Journal of Philosophy of Education*, 37(3): 415–425, <https://doi.org/10.1111/1467-9752.00336>
- Carter, I. 2003. "Positive and Negative Liberty", *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/win2019/entries/liberty-positive-negative/>
- Christman, J. 1991. "Liberalism and Individual Positive Freedom". *Ethics*, 101(2): 343–359, <https://doi.org/10.1086/293292>
- Christman, J. 2018. "Autonomy in Moral and Political Philosophy". *The Stanford Encyclopedia of Philosophy* (Spring 2018), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/spr2018/entries/autonomy-moral/>
- Hewer, C. 2001. "Schools for Muslims". *Oxford Review of Education*, 27(4): 515–527, <https://doi.org/10.1080/03054980120086211>
- Kymlicka, W. 1995. *Multicultural citizenship: A liberal theory of minority rights*. New York: Clarendon Press.
- Rawls, J. 1993. *Political Liberalism*. New York: Columbia University Press.
- Snik, G., & De Jong, J. 2002. "Why Should States Fund Denominational Schools?". *Journal of Philosophy of Education*, 36(4): 573–587, <https://doi.org/10.1111/1467-9752.00296>
- Snik, G., & De Jong, J. 2005. "Why Liberal State Funding of Denominational Schools Cannot be Unconditional: A reply to Neil Burtonwood". *Journal of Philosophy of Education*, 39(1): 113–122, <https://doi.org/10.1111/j.0309-8249.2005.00423.x>

Split Mind in a Split Time

The Disturbance of Internal Time Consciousness in Schizophrenia

Matheus Paim

Temporal experience becomes fragmented in schizophrenia. To better understand time disturbance as a symptom of schizophrenia, the present paper proposes an interpretation of the schizophrenic symptomatology in terms of Husserl's conception of internal time consciousness. Furthermore, the prevalence of time disturbance may be explained by the close relation between self and time.

1. Introduction

"While watching TV it becomes even stranger. Though I can see every scene, I don't understand the plot. Every scene jumps to the next, there is no connection. The course of time is strange, too. Time splits up and doesn't run forward anymore. There arise uncountable disparate now, now, now, all crazy and without rule or order" (Fuchs 2007, 233). This is the report of a subject suffering a psychotic episode in schizophrenia. It highlights the fragmented character of temporal experience in which events do not flow smoothly into each other.

Philosophers have inquired into the nature of time since ancient Greece. However, an account of our subjective temporal experience that is still widely accepted today only came with Augustine, who understood time as confined to the present moment and apprehended by the mental operations of the mind (Husserl 2019, 21; Mensch 2010, 23). For Augustine, it is in our minds that we measure time, in our memory as past, attention as present and expectation as future. Husserl, who pointed out several inconsistencies in Augustine's account of time consciousness, examined the structures of consciousness responsible for the constitution of our temporal experience. Whenever these structures function abnormally, the subject's temporal experience is disturbed. In the psychotic episode mentioned before, Fuchs (2007) considered a disruption of passive synthesis. A temporal structure of consciousness which integrates succeeding moments. Since it becomes disrupted, there is no coherence in the subject's temporal experience.

Through the method of phenomenological reduction, Husserl arrives at the object of his analysis of temporal consciousness, the immanent time of the flow of consciousness (Husserl 2019, 23). He aims to bring the constitution of apprehended time to light in order to clarify the a priori laws of time. To explain how we constitute time, Husserl examines how we apprehend an enduring object while it unfolds across its succeeding phases. Husserl proposes three moments constituting the originary temporal field, namely retention, primal impression, and protention (Bernet, Kern, and Marbach 1993, 103). Together, these moments unfold as a unity at any given instant such that consciousness extends beyond the now to grasp the temporal object. The key to temporal experience is the constant modification of the flow of consciousness. Impressions are modified into a continuous flow of retentions which are projected to the future as they sink further into the past. The double intentionality of consciousness marks the final step in the constitution of our temporal experience (Zahavi 2003, 170). Longitudinal intentionality accounts for the continuous flow of consciousness, merging its different phases and giving rise to self-awareness of the temporal flow. Whereas transverse intentionality accounts for the enduring temporal object. Together, they enable the perception of duration across the succession of elapsed phases of consciousness. In the absolute flow of consciousness, the whole of time-constituting phenomena takes place. Consciousness resembles a water stream, constantly flowing, with all its points connected to each other forming a unity, and just as the leaves and the fish float through the stream, so do the enduring temporal objects flow through absolute consciousness. However,

the importance of time-constituting phenomena may be taken for granted until our temporal experience becomes disturbed.

The Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-V) characterizes schizophrenia as a psychotic disorder with two or more of the following symptoms present for over a month: hallucinations, delusions, disorganized speech, grossly disorganized or catatonic behaviour and negative symptoms such as avolition (Tandon 2013). Moreover, a common aspect of schizophrenia is the fragmentation of the self (ego) during psychosis, in which thoughts and actions are perceived as coming from another self. (Bob & Mashour 2011, 1044). Although time disturbance is not included in this definition, it has been shown to be closely related to different symptoms of the condition (Fuchs 2007; Fuchs 2013).

Husserl's account of the consciousness of time, of how we come to be conscious of time, is also an account of self-consciousness, of how we form a sense of selfhood (Mensch 2010, 23). Consciousness manifests itself in a field of temporal relations, the now-moment surrounded by memories and anticipations. The integrity of temporality is essential for a coherent sense of selfhood and identity. In temporal fragmentation, each now-moment becomes extraneous to consciousness, disassociated from the previous or succeeding moments. Without a sense of temporal experience being 'my' experience, the individual's temporal orientation and identity disintegrates. The intertwined relation between self and time may elucidate the prevalence of time disturbance in schizophrenic patients.

In schizophrenia the structures responsible for our temporal experience such as the three moments of the originary temporal field (retention, impression and protention) and double intentionality function abnormally. As a result, the temporal experience of the subject becomes fragmented, leading to different symptoms such as disorganized speech and delusion.

To clarify Husserl's contribution, it is first necessary to explain the initial conditions for Husserl's analysis of time consciousness. I will next give a more extensive elaboration of his account on how temporal consciousness is constituted, presented in his lectures of *Phenomenology of Internal Time Consciousness*. Furthermore, I will explain how Husserl's theory can be applied to time disturbance in the psychopathology of schizophrenia.

2. The investigation of internal time consciousness

2.1 Husserl's Phenomenology of Internal Time Consciousness

The analysis of internal time consciousness begins with Husserl's precondition for any phenomenological analysis, the phenomenological reduction, "the complete exclusion of every assumption, stipulation or conviction concerning objective time" (Husserl 2019, 22). Husserl does not intend to say that the psychological and metaphysical nature of time does not exist but that all presuppositions and expectations about time must be set aside if time is to appear by its very nature in conscious experience. What is left from this reduction is the immanent time of the flow of consciousness (Husserl 2019, 23). Husserl means to arrive at the structures of consciousness constituting the modes of objects appearing as past, present and future.

Furthermore, Husserl inquires about the foundational level of our experience of time. As Sokolowski suggests, there are three levels of temporal structure (Sokolowski 2000, p. 130). Firstly, objective time can be measured and verified, for instance, when we say the flight to Berlin lasted two hours. Secondly, subjective time represents how mental states and experiences follow one another in our conscious life. For example, I may experience a two-hour lecture as fast or slow. Lastly, there is the consciousness of internal time, it accounts for our experience of the succession of subjective time (as with memory and expectation). It is

absolute as it does not need any level below it, representing the foundation of our experience of time. Consciousness of internal time makes the following scenario possible: after watching a lecture with a friend, we disagreed about its (subjective) duration, but once we checked our watches, we agreed on its (objective) duration. The absolute level of temporal structure, consciousness of internal time, provides the setting for the two levels mentioned above (objective and subjective). It is acknowledged by Husserl as giving the form to our temporal experience.

2.2 The temporal object

The temporal object reveals the core of Husserl's analysis, namely, duration, which is "the continuation of the same throughout the succession of other phases" (Ricoeur 2010, 26). For instance, in perceiving a bird flying in the air, the bird is constantly changing across the duration of its movement, yet I am always perceiving the same bird flying through the sky. The paradoxical nature of the temporal object, of simultaneously being different and still the same, represents something more than mere succession of phases or the sum of its moments. Another example of the temporal object is given by Husserl. When listening to a melody, it begins to sound, as I hear the second tone, the first tone is somewhat present, and the third tone is anticipated as if the melody is going towards a conclusion. The melody continues to sound and at every new moment, the previous tones seem to carry me along as I expect more to come. In each moment of its duration, I hear a melody as a continuation of the tones and not just a particular tone. As Husserl himself puts it:

"Every tone itself has a temporal extension: with the actual sounding I hear it as an ever new now. With its continued sounding, however, it has an ever new now and the tone actually preceding is changing into something past" (Husserl 2019, 43).

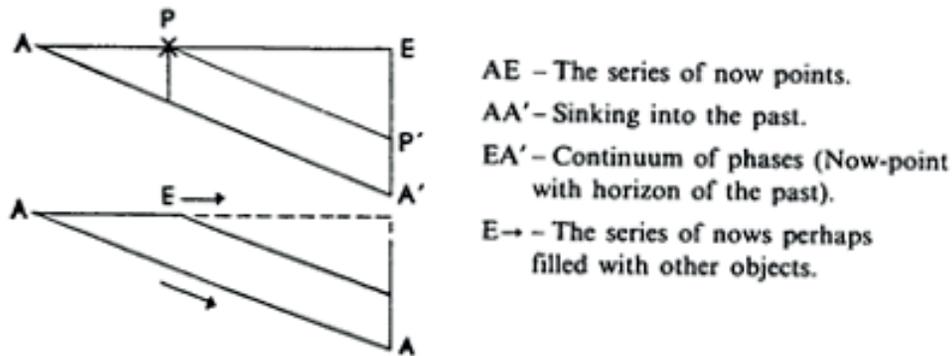
The extension of the tones in a melody is the result of the extension of consciousness beyond the now to grasp the temporal object (the melody in this case). In turn, the extension of consciousness beyond the now is a result of the 'time constituting phenomena', which represent the originary temporal field of our experience at any instant (Bernet, Kern, and Marbach 1993, 103). It is composed of three different moments (modes of appearance); retention, primal impression and protention.

Considering the above example of the melody, the tone occupying the now-moment corresponds to the primal impression. It is the actual phase of immediate experience. But almost instantly, the tone in primal impression slips into the past. However, it does not disappear from consciousness, but is now retained. Retention is arguably the key to our grasp of the temporal object, "in this sinking back, I still 'hold' it fast, have it in a 'retention', and as long as the retention persists the sound has its own temporality" (Husserl 2019, 44). The tone has changed (from actually sounding) into something 'just past', as 'having been'. As new tones sound, the previous retentions are again retained, being constantly modified in its relation to the actual now.

As the melody unfolds, the living present encompasses a chain of retentions, of fading phases belonging to the now. This process may be understood as a comet's tail, the now actual phase is the center of the comet, accompanied by its elapsed phases as its tail (Sokolowski 2000, 146)²⁰. As an ever new now is generated, the previously sounding tones accumulate and decrease in clarity. At this point, Husserl argues "as the temporal object moves into the past, it is drawn together on itself and thereby also becomes obscure" (Husserl 2019, 47). Across the succession of its phases, the object is in constant change, continually different in its passing away. Yet, its continuity is maintained, its parts are unified as the retention ceases. The remaining moment of the living present is protention, directed to the 'not yet'. It is based on retention as a 'project shadow' or a projection of the past, as an expectation, into the future (Bernet, Kern, and Marbach 1993, 103).

²⁰ In fact this analogy was first made by Husserl (Husserl 2019, 52)

Husserl designed a diagram to help illustrate the identity of the temporal object in the succession of its different phases (see fig. 1). As time flows along the horizontal line, in each instant an impression sinks down the vertical line as a retention, and the retention is further retained as the temporal object unfolds. At the end of the vertical line retention ceases and the retained phases fall into oblivion. Suppose I hear the melody do re mi, then at time E in the diagram, I have the retention do re (represented by A'P'). Although protention is not represented, at the time E with retention do re, there is also the protention mi, projected forward.



(fig. 1: Mensch 2010)

In this diagram, all moments are connected through a continuum of constant modifications in which every elapsed phase is modified as a new now appears. In immanent time, past, present and future are intertwined, unfolding in every single instant. When I hear a particular tone of a melody, I am not hearing independent tones but a melody. Thus, I am aware of the whole (melody) through its parts (tones) (Wehrle 2019, 6).

2.3 Beyond retention: Recollection

We must consider recollection, or secondary remembrance, in order to account for what is no longer present, for that which lies beyond the comet's tail (Husserl 2019, § 14). In listening to a melody like do re mi, the content which is retained in my perception is the same as when this same melody is remembered. Recollection is a 'presentification', of the original experience, itself not a perception, but a reiteration of the original perception (Bernet, Kern, and Marbach 1993, 105)²¹. After listening to a melody, I can represent (recollect) any of its present instants, which 'mirror' the continuity of retentions and protentions of the (originally) perceived temporal object. In other words, since what is given in perception is identical to what is recollected, every past retained may be arranged as a quasi-present with its own protentions and retentions (Ricoeur 2010, 32).

The key aspect to understanding recollection is the passage from perception to non-perception. Retention, composed of a continuum of gradations being constantly modified, reaches an 'ideal limit', the limit of the pure now (the end of the comet's tail). The origin of the past lies in retention, not in recollection; as Husserl claims, "For only in primary remembrance do we see what is past; only in it is the past constituted" (Husserl 2019, 64)²².

²¹ In perception the temporal object is presented, the object unfolds in its originally given mode as a presentation.

²² For Husserl, the past is directly perceived (sensed) as past. Thus, retention is itself a perception, inherent in the present moment.

2.4 Double intentionality and absolute consciousness

Husserl's inquiry aims to investigate the immanent time of the flow of consciousness. The term 'flow' is used throughout Husserl's analyses. It indicates the 'absolute subjectivity' from which the primal source point arises. Husserl admits that there is no proper name to refer to the absolute consciousness, and so he uses the term 'flow' (Husserl 2019, 100). To better understand absolute consciousness, we may imagine it as a stream, a flowing watercourse. The stream does not stop flowing, every point is constantly renewed by upcoming water and all parts of the stream are connected, unifying it. Absolute consciousness works in the same way. As long as we live, temporal objects endure in our perception, much like fish swimming along a portion of the stream, emerging and then gradually disappearing to more distant waters. To know exactly what is the true nature of the stream would be equivalent to the fish being aware of the water which surrounds it or the bird being aware of the air in which it flies. Absolute consciousness is the absolute ground of all being, it is the condition for the possibility of our temporal experience.

Longitudinal intentionality runs through the flux in which primal impression changes to a retention and then into a retention of a retention (Husserl 2019, 107). This intentionality enables the individual to 'knit together' an impression with its retentions and protentions and it makes awareness of the temporal flow possible (Russell 2006, 134). In the example of the water stream, it is because of this intentionality that I perceive a fish continuously swimming downstream as my longitudinal intentionality connects all its temporally appearing phases. It is the fact that 'I' am aware of my own streaming consciousness, a 'pre-reflective self-manifestation of consciousness' which enables temporal self-awareness (Zahavi 2003, 171). In transverse intentionality on the other hand, reflection is turned from the primal sensation to the enduring object, and thus perceived as a temporal object (Dyke and Bardon 2013, 402). By means of this intentionality, I am able to intend a sound over its successive appearing phases. In the example of the water stream, it is because of transverse intentionality that I see the same fish throughout the duration of its swimming.

Transverse intentionality is responsible for the departure from the living present to the awareness of the temporal object and longitudinal intentionality relates and merges the different phases of the appearing temporal object. In the example of the water stream, constituted immanent time is the fish and leaves which flow along the stream, corresponding to temporal objects enduring in consciousness. Pre-immanent temporality is not the water stream itself (that is absolute consciousness) but the water currents which direct and guide the flowing content along the stream. The two intentionalities are interwoven, for the existence of a temporal object, temporal phases of its appearing are necessary, and for the existence of a temporal order through appearing phases, a relation to the enduring temporal object is necessary (Russell 2006, 134).

In understanding the flow, as absolute subjectivity, Husserl claims "we can only say that this flux is something in conformity with what is constituted, but it is nothing temporally 'Objective'" (Husserl 2019, 100). As a result, temporally constitutive phenomena are themselves non-temporal, retention does not itself happen in the now, and as such we cannot use temporal predicates to the phenomena which constitute time. This is the essential difference between the constituting and the constituted (Mensch 2010, 90). If the time constituting consciousness was itself within time then we need another consciousness in order to constitute the time consciousness as temporal and thus would result in an infinite regress (Bernet, Kern, and Marbach 1993, 108). That is to say, although temporal objects have their objective temporality, that which constitutes them does not. I may be able to say how long it takes for the fish to swim or for the leaves to float down the stream but I cannot say the same for the stream itself, neither its water currents. The stream and the currents make up what we come to understand as past, present and future. By means of the absolute consciousness, time in its appearance unfolds, making possible the constitution of the immanent temporal unity and the temporal objects.

3. Internal time consciousness and Psychopathology of temporality

In the field of psychopathology, the study of the disruption of temporal experience encompasses conditions such as obsessive-compulsive disorder, depression and schizophrenia (Fuchs 2013). In this regard, Husserl's investigation of internal time consciousness proved to be a fruitful source for a conceptual understanding of the psychopathology of temporality (Vogele, Kai, and Kupke 2007; Stanghellini 2016; Fuchs 2007; Fuchs 2013).

In the previous sections, it was shown how time-consciousness guarantees our perception of temporal objects as well as our awareness that 'I' am the one experiencing the object over time. Without the structures of the living present, namely, impression, retention and protention, I would never be able to make a statement like: "The bird has been hovering in the sky for several seconds until it rapidly dived into the water to get a fish, I felt vibrant throughout the occasion". This example not only presupposes the perception of an object as enduring and having a temporal location, but it also presupposes a pre-reflective self-awareness, that 'I' am the one feeling vibrant throughout the succession of the object. In the next section, I will elaborate on how Husserl's analysis of time-consciousness may support an understanding of the psychopathology of temporality. I will narrow the investigation to time disturbance in schizophrenia.

3.1 Schizophrenia and time disturbance.

Schizophrenia is a psychotic disorder characterised by hallucinations, delusions, disorganized speech, grossly disorganized or catatonic behaviour and negative symptoms (such as avolition) (Tandon 2013). These symptoms significantly impair the person's social relationships as well as their occupations. It has been suggested that these symptoms may be explained by a disturbance of internal time, by a disruption of the constitution of time consciousness and the pre-reflective self-awareness of our temporal experience (Fuchs 2013). However, we may first ask ourselves if Husserl himself saw the possibility of an error in our temporal constitution. In section 22 of *Phenomenology of Internal Time Consciousness*, Husserl states "The individual apprehension can be wrong, corresponding to no reality" (Husserl 2019, 73). That is to say, I may indeed experience a temporal succession but it does not mean that I have experienced an objective event in the world. The mismatch between my individual temporal succession and an external (objective) happening relates to the phenomenon of hallucination in schizophrenia in which people see and hear things that others do not. Although Husserl admits a possible discrepancy between temporal succession and the correspondence to a transcendent object. It may not account for a disturbance in the synthesis of time-consciousness in which a temporal succession barely (or does not) takes place. In this regard, we move to a consideration of authors drawing on Husserl's conception of internal time consciousness.

In discussing time disturbance in schizophrenia, patients seem to suffer from a disruption of longitudinal intentionality. By means of the pre-reflective awareness of the flow of consciousness, a retention of the elapsed phase of a temporal object acquires the sense of being 'my' experience (Zahavi 2003, 171)²³. A disturbance of this phenomenon may explain different symptoms in schizophrenia. For example, it is common that individuals cannot keep up a conversation. They lose track of their line of thought and their speech becomes fragmented (Wehrle 2019, 14). Consider the following report of a patient: "I can concentrate quite well on what people are saying if they talk simply. It's when they go into long sentences that I lose the meanings. It just becomes a lot of words that I would need to string together to make sense." (Fuchs 2007, 233).

²³ In each moment there is an awareness that it is 'my' experience which is passing away.

In this situation, it seems that longitudinal intentionality, which ‘knits together’ protentions and retentions to impressions, is impaired and the spoken sentence which unfolds as an enduring (temporal) object disintegrates. In addition, it may be argued that since there is a link between time constituting phenomena at the originary stage and what is remembered in recollection, the effects of the disruption should also manifest in reproduction (memory)²⁴. This is indeed the case, as research on “autooetic awareness” (to relive one’s past events) and memory recognition shows a disturbance among schizophrenic patients, in which they cannot connect separate events into a memorable coherent whole (Vogeley, Kai, and Kupke 2007, 162).

Fuchs (2013) emphasized the impairment of protentional functioning in schizophrenia. Protention involves a ‘cone of probability’ in which consciousness is directed towards more probable events as determined by retentions and impressions, thereby avoiding inadequate associations and distractions (Fuchs 2013, 85). Let us consider the following description of a schizophrenic subject: “When I move quickly it’s a strain on me. Things go too quickly for my mind. They get blurred and it’s like being blind. It’s as if you were seeing a picture one moment and another picture the next.” (Fuchs 2013, 86). According to Fuchs (2013), with the disruption of protention, events happen too fast to be consciously perceived. As a consequence, the individual is unable to intend a future course of action, and becomes confused.

In another example, a patient reported “I could no more think what I wanted; constantly alien thoughts were pushing in between [...] as if someone would not think by himself and would be prevented from his own thinking, and his thoughts were controlled [...] I began to wonder whether this was still me or an exchanged person” (Fuchs 2007, 233). In terms of time disturbance, this experience may be interpreted in the two following ways. Firstly, the example may illustrate a fragmentation of the pre-reflective awareness of consciousness, in which temporal flow of consciousness is aware of its own streaming (Zahavi 2003, 170). The individual may experience the thoughts as coming from another person precisely because he lacks the self-awareness of the temporal flow which prevents the subject from becoming disassociated from his own mental experience as it unfolds in time. In other words, objects are enduring and succeeding each other but there is no ‘I’ which they belong to. Without the ‘self’, the experience is perceived as coming from another person. Secondly, Fuchs argued that since protention is disturbed, the individual is no longer able to exclude inadequate associations and thoughts, in a way that ‘consciousness is surprised by itself’²⁵. Without an active orientation towards the future, one is forced to concentrate on the moments which just elapsed from consciousness (Fuchs 2013, 86).

Moreover, there are also specific impairments in the retentional aspect of consciousness. A study reported that schizophrenic patients often experience a *déjà vu*, meaning events are experienced as if they already happened before. For example, a subject reported “When I heard the news, I felt I had heard it before” (Stanghellini 2016, 50). This indicates a fragmentation of time in which retention cannot be distinguished from impression. As a consequence, the individual feels that the actual moment already occurred and the present is perceived as being past.

In normal circumstances, when we are completely engaged in our present activities, we seem to become unaware of our bodies and even time itself. Our whole being is carried along as we perform the action at hand, as if we are ‘inside time’. Whenever this implicit temporality is disrupted, one goes out of synchrony

²⁴ Reproduction is the re-presentation of an earlier perception of a temporal object. If there is a problem in the original perception, this should necessarily reflect on the reiteration of the original perception, hence in the reproduction.

²⁵ The explanation of Fuchs (2013) becomes clearer in terms of the “cone of probability” mentioned above.

with the 'intersubjective now' of others (Wehrle 2019)²⁶. The experience of time becomes explicit and the individuals' experience becomes faster or slower in relation to others. In schizophrenia, individuals may experience a 'delusional mood' so that ideas and thoughts have to be realised as soon as they come to mind, hence the subject becomes startled and impatient. As a result of this delusion, they feel uncomfortable in the presence of others as they are not sharing a common time. This in turn, may explain the isolation and avolition which schizophrenic individuals experience (Fuchs 2005, 197). In addition, the lack of synchrony with the time perception of others may explain the findings of empirical studies on time estimation tasks. It has been shown that when schizophrenic patients are asked to estimate the duration of an interval, they tend to overestimate as well as underestimate its time, judging the presented stimulus as faster or slower than its actual duration (Vogeley, Kai, and Kupke 2007, 161).

Husserl identifies the constitution of time with the constitution of the self (Mensch 2010, 3). Hence, Husserl's analysis of time-consciousness may be understood as an account of self-consciousness. As Zahavi (2003) argued, the retentive modification not only enables the experience of an enduring object through its successive phases, but it also provides a "pre-reflective and inherent temporal self-awareness". In the structure of time-consciousness there is an intrinsic pre-reflective awareness which makes possible the perception of the self as well as the temporal object, hence time and self may be understood as inseparable.

Self-consciousness is to be aware of oneself in the passage of time. Without the self, we are just consciousness, a flow of events and experiences without any structure. The distinction between self and the consciousness of which it is conscious, is equivalent to time constituting phenomena and the constituted temporal objects.

Before acquiring an objective temporality (with identity and duration), temporal objects exist in the flow of absolute consciousness. It is as if we give 'form' to the absolute flow by means of our temporal structure. Hence, we create a temporal order in which things have a beginning and an end. Time and self are equivalent to the extent that both may be understood as structures, 'forms' in which we give order to reality. Thus, on the one hand we have consciousness and time constituting phenomena as the ground of all being, the 'matter' or content of reality. On the other hand, there is the 'self' and the constituted temporal objects as the structure, the 'form' of which we order the content.

In summary, schizophrenic patients often report intruding thoughts and the feeling that these thoughts belong to a different self, this relates to a disintegration of both the protentional function to direct one's future actions as well as the synthesis of temporal constituting phenomena (Vogeley, Kai, and Kupke 2007). Given the intertwined relation between the constitution of the self and that of time, it is expected that the fragmentation of temporal experience occurs together with fragmentation of the self. This may explain the prevalence of time disturbance in different symptoms of schizophrenia.

4. Conclusion

Time is fundamental to our lives, it is familiar and yet it is remote, quite unknown to us. There have been several philosophical accounts on the nature of time since ancient Greece. Husserl investigates the structures of consciousness which makes our temporal experience possible. By means of phenomenological reduction, he derives the immanent time of the flow of consciousness, which is the object of his inquiry.

²⁶ As Wehrle (2019) noted, there are also positive experiences of explicit time (for example, explicit time may be a necessary feature of human embodiment). But since we are considering schizophrenia, explicit time is looked upon negatively, as a burden to everyday performance.

Husserl examines the enduring temporal and arrives at the three moments constituting the originary temporal field, namely retention, primal impression and protention. At any given instant, the three moments unfold as a unity, allowing consciousness to stretch beyond the now to grasp the temporal object. The key element in our temporal experience is the constant modification of the continuous flow of impressions into retentions, which are projected to the future as they fall into the oblivion of the past. Furthermore, there is the operation of the double intentionality of consciousness. It consists of, on the one hand, longitudinal intentionality, which connects all different phases in the flow of consciousness, and it also produces a self-awareness of this flow. On the other hand, there is transverse intentionality, which gives rise to the enduring temporal object. The structures of inner time unfold in the flow of absolute consciousness to which the whole of time constituting phenomena takes place. This flow may be better understood as a water stream with fish and other beings navigating across it.

Husserl's analysis of internal time consciousness gives an account of how time and its objects are constituted. Temporal experience forms the basis of the operations of consciousness. However, the constitution of time and its objects may be disrupted on certain occasions. In schizophrenia, the subjective experience of time is disturbed. Husserl's investigation of internal time consciousness has been applied to the disturbed temporality inherent in various symptoms of schizophrenia, such as delusions, disorganized speech and avolition. The individual is unable to plan for future actions and there is a general lack of agency. The present work has shown how some of the symptoms of schizophrenia may be better understood in terms of a disruption of protention, retention and longitudinal intentionality.

To conclude, Husserl offers an influential account of our temporal experience. When applied to schizophrenia, this account may well explain the impaired temporal experience as a structural disturbance of time consciousness. As argued above, the close relation between the constitution of self and time may explain the prevalence of time disturbance in different symptoms of schizophrenia. Yet, both the fragmentation of the experience of self as well as the fragmentation of temporal experience occur concurrently in schizophrenia. Thus, it is plausible to postulate that similar, if not identical, structures constitute self and time, and that once the underlying structure is disrupted, the experience of both becomes fragmented. Nonetheless, it may be advised to further examine Husserl's account in order to better understand the pathological impairment of temporal experience, and to draw more concrete conclusions which will help to carry out further empirical studies.

References

- Bernet, Rudolf, Iso Kern, and Eduard Marbach. *Introduction to Husserlian phenomenology*. Northwestern University Press, 1993.
- Bob, P., & Mashour, G. A. (2011). Schizophrenia, dissociation, and consciousness. *Consciousness and Cognition*, 20(4), 1042-1049.
- Dyke, Heather, and Adrian Bardon, eds. *A Companion to the Philosophy of Time*. Malden: Wiley-Blackwell, 2013.
- Fuchs, T. (2005). Implicit and explicit temporality. *Philosophy, Psychiatry, & Psychology*, 12(3), 195-198.
- Fuchs, Thomas. "Temporality and psychopathology." *Phenomenology and the cognitive sciences* 12, no. 1 (2013): 75-104.
- Fuchs, Thomas. "The temporal structure of intentionality and its disturbance in schizophrenia." *Psychopathology* 40, no. 4 (2007): 229-235.
- Husserl, E. (2019). *The phenomenology of internal time-consciousness*. Indiana University Press.
- Mensch, J. (2010). *Husserl's Account of Our Consciousness of Time*.
- Ricoeur, Paul. *Time and narrative*. Vol. 3. University of Chicago Press, 2010.
- Russell, Matheson. *Husserl: A guide for the perplexed*. Bloomsbury Publishing, 2006.
- Sokolowski, Robert. *Introduction to phenomenology*. Cambridge university press, 2000.
- Stanghellini, Giovanni, Massimo Ballerini, Simona Presenza, Milena Mancini, Andrea Raballo, Stefano Blasi, and John Cutting. "Psychopathology of lived time: abnormal time experience in persons with schizophrenia." *Schizophrenia bulletin* 42, no. 1 (2016): 45-55.
- Tandon, Rajiv, Wolfgang Gaebel, Deanna M. Barch, Juan Bustillo, Raquel E. Gur, Stephan Heckers, Dolores Malaspina et al. "Definition and description of schizophrenia in the DSM-5." *Schizophrenia research* 150, no. 1 (2013): 3-10.
- Vogeley, Kai, and Christian Kupke. "Disturbances of time consciousness from a phenomenological and a neuroscientific perspective." *Schizophrenia Bulletin* 33, no. 1 (2007): 157-165.
- Wehrle, Maren. "Being a body and having a body. The twofold temporality of embodied intentionality." *Phenomenology and the Cognitive Sciences* (2019): 1-23.
- Zahavi, D. (2003). Inner time-consciousness and pre-reflective self-awareness. *The new Husserl: A critical reader*, 157-180.

“Everything is philosophy!”: An Interview with Sonia de Jager.

by Ivar Frisch and Ties van Gemert

Sonia de Jager (1988) is a second-year PhD student at the Erasmus School of Philosophy. De Jager received a BFA from the Gerrit Rietveld Academy, a MA in Film and Photography, and a ReMA in Arts and Culture from Leiden University. She teaches art theory at the Willem de Kooning Academy and organizes conferences about music, art, and philosophy. Meanwhile, she also makes artworks herself. Her PhD research is about the concepts of ambiguity, limits, and reason in the rapidly expanding field of Artificial Intelligence.

Below you will find a rendition of an interview with Sonia de Jager that took place on the 10th of October around 11:00 at the Heilige Boontjes in Rotterdam. The audio was transcribed using computer software and redacted by Ivar Frisch and Ties van Gemert.

Ties²⁷: Is there enough space in the academic community for interaction between philosophers and non-philosophers?

Sonia: I think so, but if philosophers remain dogmatic, then interaction proves to be rather difficult. It is a cliché, but armchair philosophy really is a thing. Philosophers aren't always very open academics. Of course, you can only think and write, nothing wrong with that. But in a lot of cases, this makes no sense: hands-on engagement with the things you talk about should always be part of your research. I have a very interdisciplinary background, which makes me comfortable with the idea of talking to a doctor or a bicycle repair person to make adjustments to my philosophy. For me, it is obvious that you should do this. Especially if philosophy is - as we philosophers know - this huge umbrella discipline that is all-encompassing and reaches for the edges of knowledge, the limits of reasoning, the limits of thinking. In a lot of parts of academia, there still is a dogma that leads philosophers to exclude knowledge that doesn't come in book-form.

Ties: Maybe that also has to do with philosophy's identity crisis. The fact that philosophers don't really know what their task is at the moment. And as soon as you also start involving people who do have a well-developed task like a doctor, you start thinking, what do I have to add?

Sonia: Yes, I completely agree, and it's funny: in the end, it does come down to motivation and confidence. Are you convinced of the fact that you as a philosopher have something to add? You could become very self-conscious and think: "ahh, of course, they can figure it out themselves if they really think about it". But that's not the case. A lot of "hard" scientists do not often think about the use and definition of concepts. That's where philosophers come in.

There was an example I saw recently from someone in predictive processing who was talking about how in stem cell research, different scientists were using different definitions of what a "stem cell" is. It turns out, with the help of philosophers, by means of philosophical analysis, they got way better stem cell research. That's a great example of where conceptual analysis and definition can help. Philosophers should be confident in the fact that they can do that and not think "oh, no, technology and science are eating the world, so I'll look the other way and bury my head in books to stick it to the system". It all is very much related and all

²⁷ Ties van Gemert is a PhD candidate at the Tilburg Center for Moral Philosophy, Epistemology and Philosophy of Science (TiLPS). His research project "Who is Afraid of Psychology? Reconstructing and Reconsidering the Psychologism Debate" is funded by NWO (PGW.20.021).

goes hand in hand.

That's why I want to stress that public philosophy is so important. Because everyone has these questions, everyone out there is a philosopher. Everyone is interested in "what the fuck is this?" and "how the hell does that function?". So if you make philosophy more available, publicly, then people will definitely get into it. You just have to motivate them and make them feel confident about it. Because the truth is that everyone likes philosophy. It's in the shows we watch, in the books we read, in the jokes we tell, it's in all of culture. There is philosophical meaning there, urging to come out. It's not just entertainment or pure pleasure. Everything is philosophy!

Ivar²⁸: Do you think that your interest for Hegel also comes from this? That contemporary philosophy should incorporate non-academic types of philosophy?

Sonia: Yeah, for sure. I got interested in Hegel because, on a lot of roads I traveled intellectually, I always ended up bumping into Hegel. Like everywhere I went, all roads led to Hegel.

Ties: Can you give an example of that?

Sonia: When I was a kid, I was a very obsessed reader of Jorge Luis Borges, an Argentinian writer. Borges is very Hegelian. He wouldn't say that about himself, but in his work it's always about this: the thing is and is not itself, the thing is not defined by itself, meaning is a complex, ever-changing, eternally-returning construction. Identity comes from reflection, transformation, and comparison, the dialectical transform between it and the other. There is always this morphing triad. It's a loophole. You don't have a definition, something stable you might want to hang onto. You just have the evolving relationship.

So through Borges, I got to Hegel. Just by researching this idea. Of course, Hegel is hugely problematic. Yet, in a way, I think no one has been able to challenge Hegel. For me, Hegel ultimately just provided the key to what you can say about thought - without killing it in the process. He's like, look, you're always going to have that unstable shift. In any definition or in anything that you want to grasp rationally, you have that transformation effect. Things defining each other and sort of evolving because they keep answering and echoing back to each other. So that is why I think Hegel is key. But of course, he also said that women are shit. That they've got the status of "animal" and whatnot. Whacky guy. That kind of stuff. So he's also an idiot in a way, he's stuck in his time, but should not be forgiven for that.

Ties: There are a lot of objections to Hegel's system - if you perceive it as a system. How do you see these objections? Are they valid? Or do you think that these have to be "worked through"? Like Derrida would say?

Sonia: Yeah, in a lot of ways I would take the Derridean way for sure, Hegel goes on beyond Hegel. I was actually talking to Cecile Malaspina²⁹ recently about Hegel's system in this regard. Since she knows a lot about Gilbert Simondon³⁰ I asked her about the relation between Simondon and Hegel. She said that there was this huge anti-Hegelian quarter in French thought for quite a while. There was this kind of caricature, perhaps this misunderstanding, of Hegel.

²⁸ Ivar Frisch is a third year BA philosophy student at the EUR mainly interested in philosophy of technology and philosophy of mind.

²⁹ Cecile Malaspina is a French translator and philosopher. She is well-known for her translation of Gilbert Simondon's *On the Mode of Existence of Technical Objects* (1958) and a monograph on noise called *An Epistemology of Noise* (2018). Last year, she was elected as a six-year chair for the prestigious College International de Philosophie established by Jacques Derrida, François Châtelet, Jean-Pierre Faye, and Dominique Lecourt in 1983.

³⁰ Gilbert Simondon (1924-1989) was a French philosopher known for his theory of individuation in technology. His work profoundly influenced Gilles Deleuze, Bernard Stiegler, and Yuk Hui.

Ties: Like Kojève's³¹ interpretation?

Sonia: Yes, exactly. Of course, you can't say there is a "right" interpretation, but there was a very specific interpretation of Hegel and then there was a widespread reaction against this interpretation. So I think you definitely just have to "work through" these objections. But if Hegel was alive today, I don't think he would be reading Hegel, if you know what I mean. He might be reading neuroscience! With his very specific critical angle to it. Not to go "the way of the understanding" that flattens things out. But just sort of trying to go for the life of Geist and keep the fire burning, haha.

Ivar: What do you mean by that? Keep the fire burning?

Sonia: Being in between understanding and reason. The understanding wants to have clear categories and wants to have stable, unchanging definitions. But Hegel says that reason doesn't work like that. Because reason is constantly in motion, constantly developing. That's uncomfortably superrationalist, but in a way which accommodates reason to be this constantly self-tripping thing as well. Hegel would define this as reasoning perfecting itself, perhaps. But I define it for my purposes and interests like a process of constantly tripping and not falling - as if you're just rolling, unfolding downhill and you can't do anything. I'm a pretty stubborn determinist (pun intended).

Ties: How do you deal with the fact that there have been so many people who spent their lives just reading Hegel and developing interpretations of his work? If you want to say something about Hegel, it sometimes seems like a very daunting task. How do you deal with that?

Sonia: For me, that's actually one of the main reasons why I move into and away from Hegel all the time. I don't want to become one of those scholars who is like "this is my Hegel!" and then defend that interpretation. Like Pippin³² or Žižek do, for example. I think Hegel is super useful for the reasons I just mentioned. But I also think that there is a moment at which I divert from him. I understand why giving a compelling interpretation of Hegel is someone else's thing. That's great. You do your thing. But for me, I use Hegel for what I'm interested in. I use bits and pieces. For example, there's this particular aspect of reason that I think is great. We should embrace this so that we never fall into the trap of thinking that there's an objectivity, one kind of rationality that we have. For Hegel there is this "one" kind of this or that. I think I can move through Hegel, perhaps not with him. Or is that harassment?

Ties: So you don't want to develop a systematic interpretation? You just want to use him?

Sonia: Yes, as an example of a philosopher who actually shows you how science can carry itself through without falling into its own trap. In a different realm a philosopher called Nicholas Maxwell³³ is developing a challenge to how we think about scientific progress called aim-oriented empiricism - which is, in some ways, very similar to Hegel's absolutism.

31 Alexandre Kojève (1902-1968) was a Russian-born French philosopher. His seminars on Hegel's *Phenomenology of Spirit* influenced a whole generation of French philosophers including Georges Bataille, Jacques Lacan, Maurice Merleau-Ponty, and Jean-Paul Sartre.

32 Robert B. Pippin is an American philosopher. He is the Evelyn Stefansson Nef Distinguished Service Professor at the University of Chicago. His books on Hegel include *Hegel's Idealism: The Satisfactions of Self-Consciousness* (1989) and *Hegel's Realm of Shadows: Logic as Metaphysics in Hegel's Science of Logic* (2018).

33 Nicholas Maxwell is a British philosopher of science. Before retiring, he taught at the University of London for nearly thirty years. His most important books are *From Knowledge to Wisdom* (1984), *Is Science Neurotic?* (2004), and *In Praise of Natural Philosophy* (2017).

Maxwell argues that we can identify these different types of rational projects, but that in order to proceed towards what we intuitively assume as progress, we should be completely clear about the aim-oriented part. What are the ontological assumptions behind modern scientific rationality? Only then are we going to do “rational” science. Maxwell says that the way science actually functions right now is completely irrational. If we can define our goals, then we’re going to do rational science. So, in many ways, it’s similar to Hegel: define a goal before setting out to it.

It’s like home economics. You have food and you have to think: “this goes there and that goes there”. Of course, things are way more complicated than you can make them to be in a very simple cooking experiment. But, also, they’re not. You need to do the cooking experiment, you need to eat, you need to get to the next day. Sleep at night. Things are more complicated than we schematize them to be, but we need to simplify them for the sake of carrying things through - especially when you’re living in large groups. Groups are complex, but complex things are also quite simple.

Ivar: Do you mean that you need a goal to work towards, some kind of schema, but that you can never reach it? That it will always fail?

Sonia: Yes and no. You need to be aware of the fact that this goal is not the ultimate one. It’s always revisable. That’s where Hegel comes in for me. You’re not saying that there is a perfect result or some kind of fixed rationality - which is the big bias of a lot of interpretations of what ‘science’ is.

Ivar: So, let’s take a small detour and discuss one of your artworks: H-64. There is a quote in there which sounds very Hegelian: “Base what you do on that which you observe, understand observing, observe understanding and only after that observe and understand.” Could you tell us something about this artwork? Was it indeed inspired by Hegel?

Sonia: Sure. That is not a quote from Hegel, but there is one somewhere else in that work. At that time I was working on a lot of text-based artworks. So, I made this thing called H-64³⁴ (the last hexagram of the I-Ching: “not yet fording, not yet complete”), based on *The Norton Anthology Of Theory And Criticism* (2001). I just took entire quotes from the book, chronologically arranged them, from the Presocratics all the way to now, I made a new text out of all these quotes. I didn’t change anything in the quotes, yet I made a new text out of them. There’s only one quote which does not originate in the anthology but comes from Frank Zappa’s *Joe’s Garage’s* central scrutinizer.

I was constantly fighting with this idea of “originality”. What can I say is new? Language is a set of limited symbols. How can it give rise to infinity? That’s just super crazy, you are constantly making new things, while constantly going over the same old ingredients. As we know, repetition doesn’t mean that “the same thing” keeps happening. Actually, every time the repetition happens, “the new” is happening - which for me is like a magical philosophical idea.

Every time you read it, if you take the trouble to actually read it, it almost becomes this strange mantra. Like “do your best, do your best, do your best”. They are all supposed to be variations on Golden Rule type of sentences. Rules that we all know - that we’ve been repeatedly exposed to. But obviously we still don’t know these rules because things keep going wrong. We’re not following those rules.

³⁴ The work can be found here: www.glass-bead.org/research-platform/notyetfording/?lang=enview

That's why I like the insistence. Fucking copy-paste it. So that it all gets wired into the brain. But at the same time, it's about making a joke. That it's not possible - even though you have this information right in front of your eyes, to actually proceed accordingly. You're not going to do your best! You're just not gonna! You know what I mean? H-64 also presents you with this daunting task of reading the whole thing, so you can also just play with it by scrolling and going back and forth.

Ties: I think this also happens when you read a text by a philosopher. You don't read all of it, you're not reading every sentence. Sometimes you just skip some parts of a paragraph. In this sense, style is very important. Because it gives you the rhythm for reading...

Sonia: For sure. An interesting example is the following: a book is the "same" book everywhere, it is printed and distributed worldwide. But no one is ever going to read the same book. You are in your specific context and that is going to affect how you read the book. But also literally every single word that you encounter is a different word than everyone else's. You have all these different types of interpretations, and they all come from this "one" book. But it's neither what the person said, nor what you interpret. Nor is it what's in between, because what's in between is also unique in between you and whatever thing you encounter. So again, it's one of those magical philosophical things. Everything is different, everything is the same. But also not. There's no identity whatsoever. That's wonderful.

Ties: That's where hermeneutics comes in right? You start thinking about what is my horizon, and what is the text presenting to me...

Sonia: Sure. You see, that is where people make a lot of philosophical mistakes. When politicians or scientists argue, for example. If you actually deconstruct what's going on, it comes down to them having a different definition of a particular concept in that discussion. That's it. It's very easily said, but it often very much comes down to that - which is why I think philosophy is super important. Because it aims to really deal with that. To disentangle that, only to tangle up other things. To think carefully and to be careful with how you use concepts...

Ties: Would you say that's one of the primary tasks of philosophy, namely to develop concepts and clean up concepts? In that sense being analytic?

Sonia: Yeah, I wouldn't say "clean up", because then you do get very analytic...

Ties: What's wrong with that?

Sonia: What's wrong with "cleaning up"? Well, first of all you assume something is "dirty". "Cleaning" falls prey to the bias that you can get a cleaned up concept when in fact you can't. So that's where continental philosophy comes in. You never have a cleaned up concept. Fostering awareness about that is way more important than trying to get a cleaned up concept, trying to define an object or stipulate a definition that is the same for everyone. Like, no! Let's just get to the fact that we're all in this together and we're constantly hashing it out together. Which is way more fun! That's why these conversations keep happening. That is why I'm interested in doing this. Having these feedback moments with each other. It's not about the fact that afterwards you're like: "OK, here's the interview". It's all about the transformation. This expanding, this changing of knowledge over time. That's what's fun!

Ties: What's the role of the concept of truth, here? Because if you go into this discourse of limits and difference, then a lot of times the concept of truth ends up being problematic.

Sonia: Yeah, sure. I don't go with that concept. I think people can use it, but I don't use it.

Ties: So you don't use it? You say you don't need the concept of truth?

Sonia: I don't like it and I don't need it. And I think it's a very difficult concept: how would you define it or how would you use it?

Ties: I would start with a very minimal concept of truth that's not based on correspondence, but on the simple fact that it's clear to both of us that there is something available here that looks like a cup. We don't need to have the same picture of the cup in our heads - although some philosophers have thought that was a great idea. At the very least, there is a common object that we can talk about. I think from there on you can develop a coherence theory of truth. We can see how the cup functions within the logical space of reason. That's the term I would use - because I really like Sellars.

I think you can develop a concept of truth that still does justice to the discourse on difference. I think that even within analytic philosophy, a lot of philosophers from the 1960s onwards including Sellars, Quine and Davidson, they also stress the fact that a concept is indeterminate. Think of Quine's indeterminacy of translation. So, they do justice to the discourse on difference, but at the same time keep a definition of truth. A concept of truth that I think is very useful.

Sonia: Yeah, I would agree with some of what you said. Although I would say that the part I would agree with the most is the context and functionality aspect. Because the cup is only whatever you need it to be. For us, it's now an object that we refer to. We're all putting honey into it! We're talking about it because of this conversation. But other than that, there is nothing you can say is "true" to the cup. The cup will be what you need it to be. Use it to drink, for example.

Ties: But I would say that in that discourse, the function of a cup is its truth...

Sonia: But it constantly shifts and changes, right? It will go into the dishwasher and it will then serve someone else's needs... There is this Heraclitian flux to the cup. I think truth is an important concept when it plays a role that a specific context demands of it. But I think on its own, just "truth"... That's when I'm like: I don't need it. It always needs to be a couple. In the Socratic way. Knowledge and truth is always about something.

Ties: I would say that you can't escape using a concept of truth. Otherwise it would be impossible to talk to each other. You need some concept of it.

Sonia: Well, I think that if you get all the way down to physics and ask: "what is really there?". Even in physics, as we know it right now, you have this complete indeterminacy at the bottom of reality - depending on what you're measuring and what you want to get. There is this corner where philosophy and physics meet. Depending on what you want, you are going to get a different answer. You need a reason and a context for why you're talking about it in order to assess whether it is true or false. So it's always important to say why you need it and what functions as a plane or space of reason. That's where the Hegelian aim-oriented

empiricism for me comes in - minus the blind normativity. Because when you say what you want, then you can move on. And in fact, what you want is always unstable and unsayable, but at least you're going to hash it out together. That's what thinking is.

Ties: Yeah, I see that. But let's let's put it another way. In a sense, there's a limit to what we can do with this cup, right?

Sonia: No, no. I wouldn't say so. You can fire it into space, you can shatter it. There's just so many things.

Ties: But I cannot simply say to this cup: "dissolve right now!" And make it dissolve... There is a clear limit, right? In everyday life, when you bump into a door, that is a clear limit.

Sonia: Okay, in that sense. It's a different way of thinking about what "limits" are. There are limits in the sense that you cannot do "magic" stuff, like pass through a door without opening it. Yeah, sure. But with the technology we have right now, we have gotten to the magical realm of what was previously considered not possible. And, if we don't destroy what sustains us in the process, then I think we will get to the point where we can melt a spoon with our mind. So I tend to hesitate to say that there are even those types of limits.

Ivar: So, at the same time there's always a limit, but at the same time there's no limit?

Sonia: Whatever the limit is, is the awareness of it. A limit means "here and no further". And so once you know that, you actually do know how to get further. Because you know what the limit is, once you know it, you can pass it. That's, in a nutshell, what Hegel said. So a limit is not "infinity", an ideal you never reach, a limit is actually a very clear definition on how to overcome infinity or how to engage with infinity - which is fun, philosophically. It is also realistic, we constantly make technology that challenges these limits.

Ties: I would say that science is indeed able to dissolve this cup. But how it is able to do this, I think, depends heavily on what we know to be true of the cup. Scientists need to conceptualize how this cup functions to be able to dissolve the cup. Of course, I agree that truth is context-dependent. But if you can get some clarity about the context, and I think that's possible, you can start to, not define a cup, but at least say something about what is true about the cup.

Sonia: Yeah, for sure. I would not disagree with that. I'm totally on the same page. But in the way that you're phrasing it now, you still presuppose this abstract realm in which you can measure, but any measurement you will be doing is determined by your needs. Also, "scientists" who invent a certain technology need to know and measure a certain amount of stuff that matters to them. But you don't need to know anything about electronics or telecommunication in order to send your grandma a text message. Truth, in that context, is something else.

Ties: But I think you can accommodate that. Take it up into your theory of truth. I think that context-dependency can be taken up into the theory of truth.

Sonia: Yeah, but then you need to constantly redefine the concept of truth, right? What would you say is true about that cup right now?

Ties: Of course, science changes and develops and there have been different conceptualizations of this cup. But there do seem to be clear limits to what is possible with this cup and what properties you can actually attribute to this cup.

Sonia: I think we agree, but we're kind of looking at things from different angles. I don't agree about the clarity of the limits though. Let's say that in fifty years we discover that the way glass functions is actually a very specific way in which spacetime is folded onto itself. And actually glass is not something that you can say is really "there", because it's actually a folding of spacetime which means that it's ten meters below you, and that's precisely what makes "glass" glass. You experience it as being close because of your perceptual system. But in the "true" measurement that cup is ten meters below you. That also kind of redefines the cup. Then we need to reconsider it as being ten meters below or not.

Ties: Yeah, but once again, I think that you can incorporate that. I think that's an important task for the philosopher of science to actually clear that up, but also for sociologists of science or historians of science... But don't you see it as a loss for philosophy to stop talking about truth? Because it has always been one of its primary concepts...

Sonia: I think it's something that we need to move on from, at least for now. Maybe it comes back in the future and will turn out to be very useful. But I think in the revolution that philosophy is in right now in my head, truth is unimportant. Maybe for someone else's research, it is actually very interesting. But from what I've read during my time on this planet, the philosophy of truth, I'm like, well, that's not helping us out now. But it could be useful for sure! Everything is useful, nothing goes to waste, but personally I would advise against using it and advice to study something else.

Ties: But truth has also become a thoroughly political concept, right? How do you see the relation between truth and the commons? Would you also use this way of talking and these kinds of arguments in the political realm? Do you think it is emancipating to stop talking about the commons or about truth in political discourse as well?

Sonia: Yeah, I think it's difficult. There's only so much you can do, in a deterministic world! In my research, I focus on other things. For example, with this conference Regenerative Feedback³⁵, the idea is that however disparate our practices: we all do it together. That is why this is a conference in which all sorts of walks of life are there trying to hash something out together. So it's not something specifically for philosophers. It's philosophers like Reza Negarestani³⁶, but also musicians like Dreamcrusher³⁷ and they're all in the same room together trying to figure something out. Pretty cool. So my motivation comes from the conviction that we need to mix up different disciplines, and supposed non-disciplines, so that we understand what is common, but also not common. To understand how ideas, attitudes, and methodologies differ. I consider this to be the creation of a common. So, there are completely different truths, if you will, and completely different approaches and political stances here. But then you all get into the same space in order to create the common.

35 Regenerative Feedback is Sonia's yearly conference on music and philosophy, where people like Dreamcrusher, Sjoerd van Tuinen, Cecile Malaspina and Reza Negarestani join forces to talk about stuff.

36 Reza Negarestani is an Iranian philosopher. He is one of the pioneers of theory fiction and is currently the director of the Critical Philosophy programme of The New Centre for Research and Practice. His books include *Cyclonopedia: Complicity with Anonymous Materials* (2008) and *Intelligence and Spirit* (2018).

37 Dreamcrusher is a Brooklyn-based industrial, noise, punk, shitgaze, hardcore artist working under the concept of NIHILIST QUEER REVOLT MUSIK.

The commonality is sharing, having a discussion, or creating a composition. And that's what is glorious about living together on this planet, that we can do that! Instead of thinking "ehhh... they don't know what I'm talking about". Open up that space, create the common, even if it's uncommon!

It's still a very complicated question. In the end it's all about conviction, motivation, and communication. It is all about dialogue, and you need to want to talk for there to be anything like a dialogue. Wanting to talk means you believe, to a larger or lesser degree, what you want to communicate. It's the biggest cliché thing to say. But I still believe we don't do that enough, so I will keep repeating it. Everyone tends to stay in their own bubble convinced of the ways they schematize the universe. Or they say: "fuck those people, I'm not even interested in that". But why not? We're all sharing the same space together, so we might as well try to hash it out. So that's what I would answer. You need to create the common all the time, you need to constantly work on it. And of course: it's difficult.