INSTITUTTETS SKRIFTSERIE Nr. 33
2020

INSTITUTT FOR FILOSOFI OG FØRSTESEMESTERSTUDIER

Universitetet i Bergen

Instituttets skriftserie nr. 33, 2020

Universitetet i Bergen
Trykk: Kopibutikken Bergen AS

ISBN: 978-82-90809-52-7

ÅRBOK 2020

Bergensnettverket for kvinner i filosofi

Jasmin Trächtler, Gro Rørstadbotten & Špela Vidmar (red.)

Innholdsfortegnelse

Forord

Kjære medlemmer,

Vår-semesteret 2020 går mot slutten og vil sannsynligvis bli husket av mange som svært uvanlig. Restriksjonene som fulgte etter utbruddet av Korona-pandemien endret våre liv på en drastisk og dramatisk måte. De hadde konsekvenser for våre hverdagsliv, både hjemme, på arbeid og sosialt. Og, kjære medlemmer, selvsagt påvirket også dette oss; dessverre kunne vi ikke møtes så ofte til våre interessante diskusjoner på seminarene eller til sosiale sammenkomster på lunsjmøtene.

Derfor er det en stor glede for meg å avslutte dette uvanlige semesteret med publisering av Årboken 2020. På denne måten revitaliseres også en tradisjon i BNKF. Da nettverket ble grunnlagt for over ti år siden, var intensjonen at årboken skulle dokumentere nettverkets aktiviteter og publisere innleggene som ble fremlagt på seminarene. I denne årboken publiserer vi tre seminarinnlegg og et bidrag fra masterstudent-workshopen. De tre seminarbidragene ble for første gang peer reviewed, mens det fjerde bidraget fikk grundige tilbakemeldinger på workshopen Jeg vil takke de stimulerende forfatterne Tina Firing, Ragnhild Hogstad Johrdal, Nadia Mehdi og Caroline Stampone samt våre fagfeller Steinar Bøyum, Sofie Lekve og Jesse Tomalty.

Jeg vil også takke mine samarbeidspartnere Gro Rørstadbotten og Vibeke Andrea Tellmann, samt vår assistent Špela Vidmar og nettansvarlig Vigdis Kvam. Uten disse medarbeiderne ville

ikke organisering, arrangering av lunsjmøtene, seminarene og årboken vært mulig. Tusen takk for hardt arbeid.

Sist, men ikke minst, takk til alle medlemmene for deltakelse i våre diskusjoner, for å dele tanker og ideer, og for hyggelige og sosiale sammenkomster som det forhåpentligvis vil kunne være mulig å gjennomføre i nær fremtid.

De beste hilsener fra Jasmin Trächtler

Bergen 9. juni 2020

BNKF seminarer

BNKF seminarer høstsemesteret 2019

BNKF har en lang tradisjon for å organisere seminarer med innlegg gitt av nettverkets medlemmer. Målet med seminarene har vært å støtte kvinner i deres filosofiske arbeid, og å opprettholde et positivt miljø hvor studenter kan finne kvinnelige rollemodeller. I tillegg er seminarene åpne, og alle interesserte er velkomne.

Vårt første seminar i høstsemesteret, var det en glede for BNKF å igjen kunne introdusere **Anita Leirfall**. Hennes innlegg hadde tittelen *Causal Powers in Nature and in Mind – A Kantian Approach*. Her argumenterte hun for at der er kausale krefter, som ifølge Kant, på den ene siden arbeider i naturen som ytre objekters bevegende krefter; disse forårsaker endringer i romlige relasjoner. På den andre siden, er kausale krefter til stede i relasjonen mellom objekt og subjekt gjennom affeksjon, det vil si, kausale krefter er et resultat av måten våre sinn blir affektert av empiriske objekter.

Det andre seminaret var et initiativ for et fremtidig samarbeid mellom BNKF og Kunsthallen: ved å tilby Landmark som forelesingsrom, var Kunsthallen, i samarbeid med BNKF, vert da **Tina Firing** presenterte sitt innlegg *Philosophy Without Intuitions? A Reply to the Relocation Problem*. Dette innlegget var en del av hennes masteroppgave og utdypet intuisjoners metodologiske rolle innenfor filosofi, mer spesifikt, spørsmålet om *hvordan* og *om* intuisjoner kan gi filosofisk begrunnelse, og på hvilken måte intuisjoner kan fungere som en kilde til kunnskap.

Firing stilte konkret spørsmål ved den underliggende forutsetningen om at intuisjoner spiller en sentral evident rolle i filosofi over hodet.

Det siste seminaret utgjorde en svært treffende semesteravslutning med innlegget gitt av **Caroline Stampone** hvor hun tok for seg *The Relevance of Arendt's Accounts of Evil and Refugeehood Today*. I sitt innlegg minnet Stampone oss på viktigheten av Arendts fokus på flyktningers eget perspektiv og Arendts forståelse av at ondskap innebærer å gjøre noen til «overflødige personer». På denne bakgrunnen argumenterte Stampone for at i den grad vår tids flyktningeleirer bidrar til menneskene som må leve sitt liv her, både er maktesløse og rettsløse, så er dette den sentrale situasjonen som må adresseres i dag. Hun gjorde det videre klart at med Arendt er ikke flyktninger kun en sak for statlige anliggender, men berører også individenes moralske ansvar.

BNKF seminarer vårsemesteret 2020

På vårt første seminar våren 2020 presenterte **Ragnhild Hogstad Johrdal** sitt innlegg om temaet *Constructed Kinds that Reflect Real Categories: On Hacking's Looping Kinds and His Use of Autism as an Example*. Her undersøkte hun hvorvidt det er mulig å bruke terminologi hentet fra den filosofiske diskusjonen om naturlige typer (*natural kinds*) til å snakke om klassifiseringer internt i arten menneske. Gjennom autisme som eksempel, diskuterte Johrdal hvordan klassifiseringer applisert på bevisste vesener kan føre til «looping» effekter; det vil si at individer plassert i visse kategorier blir på en måte endret gjennom kategorisering, og disse endringene vil i sin tur føre til at selve kategorien endres. Dette reiser spørsmålet at uansett om vi fortsatt snakker om en genuin type vil kategorien selv endres kontinuerlig. Johrdal valgte et realistisk syn på menneskelige

typer (*human kinds*) i den grad de fortsatt refererer til noe reelt, selv om kategoriene endres.

Vårt andre seminar denne våren var også vårt andre på Landmark i samarbeid med Kunsthallen. Med en meters avstand, var Kunsthallen denne gangen, sammen med BNKF, vert for **Nivedita Gangopadhyay** hvor hun presenterte innlegget *What is a Text? A Case Study of Joint Action.* I sitt paper, som var skrevet sammen med Alois Pichler, foreslo hun at å skape tekster, det vil dokumenter med mening, er en form for felles handling (*joint action*). Hun argumenterte for at tekster er nødvendigvis felles handling i den forstand at de krever minst to agentroller: forfatteren og leseren. Selv om tekster mangler noen av de sentrale trekkene vi tillegger hverdagslige felles handlinger, resulterer de like fullt i en sterkt og unik type felles handlingsformidling (*joint action agency*) som allerede er til stede i den skapende teksthandlingen generelt, og blir videre utvidet gjennom skapende tekster i digitale medier. Gangopadhyay foreslo i sitt innlegg at en slik unik type felles handlingsformidling har en transformativ effekt på vår sansefenomenologi når det gjelder formidling og subjektivitet.

10-årsjubileum

Hvorfor ble Bergensnettverket for kvinner i filosofi opprettet? Initiativtaker til nettverket var Mette Kristine Hansen. Hun påpekte først og fremst at det var behov for tiltak av to grunner: Først, den svært lave kvinneandelen innen filosofifaget, kvinnene utgjorde en markant faglig minoritet på alle nivåer. Blant vitenskapelige ansatte i fast stilling på fagstudiet, utgjorde kvinneandelen i 2009 kun 13 %. Så det kunne trygt fastslås at filosofi uten tvil var – og muligens fortsatt er, et svært mannsdominert fagfelt. I tillegg viste erfaring at menn har flere uformelle faglig-sosiale fora som kvinner av ulike årsaker opplever seg utestengt fra. Det innebær at kvinner kan oppleve faglig ensomhet.

På denne bakgrunn ble nettverket opprettet, og dets sentrale oppgave skulle være å bidra til å støtte opp om den faglige aktiviteten til de kvinnene som allerede har en faglig posisjon, men også ha som mål å arbeide med, og å øke rekruttering av kvinner til faget. Derfor var også intensjonen at nettverket skulle fungere som et sosialt forum der kvinner kan oppleve det å få utveksle faglige erfaringer.

Nettverket ble stiftet 21. november 2009, under en samling bestående av 34 kvinner med tilknytning til filosofimiljøet i Bergen (fast og midlertidig ansatte ved Institutt for filosofi og førstesemesterstudier, tidligere og nåværende hovedfags- og masterstudenter, tidligere og nåværende stipendiater og doktorkandidater). Arrangementet ble avholdt på Westland hotell, og var

finansiert av Universitets- og høgskolerådet og av UiB, ved likestillingsrådgiver. Vår daværende instituttstyrer, Vigdis Songe-Møller, var en uvurderlig støttespiller (og oppfølger) de første årene.

BNKFs ledere og nestledere gjennom de 10 årene

21. november 2009 ble Mette Kristine Hansen og Kristin Sampson valgt som lederduo. Sampson måtte etter kort tid trekke seg på grunn av lange utenlandsopphold og Gro Rørstadbotten overtok for henne 27. november. Fra august 2010 og frem til 30. september 2012 var Rørstadbotten leder alene. Helle Nyvold vikarierte for henne 1. oktober til 31. desember 2012. I januar 2013 overtok Kristin Sampson som leder, og Anne Granberg gikk inn som vara. Granberg overtok som leder i oktober 2014, og Paola de Cuzzani gikk da inn som ny vara. I mai 2017 fikk nettverket igjen en lederduo, Jesse Tomalty og Vibeke Andrea Tellmann. I august 2019 overtok Jasmin Trächtler ledervervet mens Tellmann og Rørstadbotten er som fast ansatte ved FoF medarbeidere i ledelsen. I tillegg er Špela Vidmar assistent.

I perioden november 2009 til mai 2017 fungerte Vigdis Kvam som sekretær for nettverket og hadde da ansvar for økonomisk oversikt, rapport/referatskriving, e-post-/medlemsliste, påmeldinger til konferanser samt ansvarlig for BNKFs nettside. Ettersom nettverket etter hvert har blitt en mer integrert del av FoF har sekretærfunksjonen blitt overflødig.

Nettverket har arrangert en rekke seminarer i tillegg til årlige konferanser. Det har også vært arrangert konferanser i samarbeid med det danske *Netværk for kvinder i filosofi*. I tillegg ble det i årene 2010 - 2013 publisert årbøker basert på innleggene som ble gitt på BNKFs seminarer. Alt dette, i tillegg til en oversikt over statutter m.m. er publisert på BNKFs nettsider uib.no/bnkf. Her legges det i tillegg ut oversikt over nettverkets

og nettverksmedlemmenes kommende aktiviteter. Nettverket har også en egen Facebokside hvor bl.a. alle arrangementer blir annonsert.

Lunsjmøter er en tradisjon som ble innført av Tomalty og Tellmann i deres lederperiode. Disse inngår nå som en regelmessig aktivitet i nettverket. Intensjonen er å gi et uformelt, sosialt tilbud hvor nettverkets medlemmer kan bli bedre kjent, utveksle ideer og dele erfaringer. Derfor inviteres alle medlemmer og kvinner som studerer filosofi til å delta.

4. desember 2019 arrangerte nettverket 10-års jubileum der det ble servert snacks og godt drikke. Arrangementet var, i likhet med alle seminarene, åpent for alle interesserte. Avslutningsvis har nettverkets nåværende leder, Jasmin Trächtler, gitt en oppsummering av denne høstens seminarinnlegg.

Øverst fra venstre: Mette Kristine Hansen, Gro Rørstadbotten og Anne Granberg. Nederst fra venstre: Kristin Sampson, Jesse Tomalty, Vibeke Tellmann og Jasmin Trächtler.

Masterstudent workshop og symposium

29-30 april 2020 arrangerte BNKF sin andre masterstudent workshop og symposium. Denne gang utforsket vi emnet "annethet" (otherness) slik det forholdet seg til kunst, estetikk og kunstfilosofi. Det ble presentert papers om litteratur, arkitektur og fotografi, hvor «den andre» fremstod som politisk, estetisk, kulturell, formal estetisk samt metodologisk og kjønnsbasert annethet. Våre hovedtalere var Danièle Moyal-Sharrock (Universitet i Hertfordshire) og Sharon Rider (Universitet i Uppsala).

Opprinnelig var arrangementet planlagt å skulle inneha tre ulike sesjoner. Første sesjon var intendert å skulle bestå av workshoper med diskusjoner i mindre grupper hvor deltakerne i de enkelte gruppene diskuterte innlegg som var sirkulert på forhånd. Alle innleggene var skrevet av master- eller PhD-studenter og veiledet av våre hovedtalere. Andre sesjon hadde form som et symposium hvor hovedtalerne presenterte sine innlegg, samt de av deltakerne som ønsket å presentere sine papers. Tredje sesjon var en bredt anlagt paneldebatt med temaet «annethet». Denne var tenkt skulle utgjøre en kollaborasjon mellom workshop-deltakerne og avgangsutstillingen – i Kunsthallen – til masterstudentene ved kunstakademiet i Bergen. I denne paneldebatten skulle også fagpersoner fra begge disipliner delta.

På grunn av Korona-pandemien og den påfølgende restriksjoner, måtte arrangementet finne en ny form. Vi arrangerte det

virtuelt med et mindre omfattende program bestående av workshop-sesjonene og innleggene til hovedtalerne. Til tross for disse uvanlige omstendighetene, var både deltakerne og hovedtalerne entydig positive i sine tilbakemeldinger angående arrangementet. De delte arrangørenes inntrykk av at innleggene og diskusjonene i workshopene var interessante, produktive og stimulerende.

Vert for arrangementet var Bergensnettverket for kvinner i filosofi. Det var organisert av nåværende leder Jasmin Trächtler, samt Carlota Salvador Megias og Špela Vidmar. Francesca Wiegand stod for konferanseplakaten og web design.

Deltakerne var Erika Brandl, Delaram Housseinioun, Nadia Mehdi, Caroline Stampone, Deva Waal og Keren Yehezkel.
BNKF takker alle deltakere, hovedtalere, organisatorer og publikum for et fantastisk arrangement.

Tina Firing

Thought Experiments Without Intuitions?[1]

Like with actual experiments, the aim of a thought experiment is to teach us something new about the world. Unlike an actual experiment, however, thought experimentation does not require interaction with the world. We can perform a thought experiment from the armchair. This might strike us as odd. If our aim is to learn about the actual world, it seems we ought to produce actual results, not imagined ones. How can the considering of an imaginary scenario lead to new knowledge of the world? Where does the new knowledge come from?

One widely accepted answer is this: thought experiments put us in a position to acquire new knowledge by inviting us to counsel our *intuitions* about the case in question. On this account, the role of intuition in thought experiments is assumed to be equivalent to the role of perception in scientific experiments (cf. Booth and Rowbottom 2014, 119). Whereas perception provides us with new information by making us aware of concrete reality, intuition provides us with new information by making us aware of abstract reality.[2]

[1] The material that makes up this contribution is part of my master's thesis. The thesis will be made available at https://www.duo.uio.no/ fall 2020.

[2] See Chudnoff (2013), for instance, for this view.

This reply to the *Epistemological Puzzle of Thought Experimentation* is, I think, false. Thought experiments are not "telescopes into the abstract realm." (Brown 2004, 1131) My primary aim in this essay will not, however, be to reject the intuition-centered account of thought experiments.[3] My hope is instead to provide a plausible intuition-free account of how thought experimentation gives rise to knowledge. More specifically, I shall attempt to circumvent an argument that I will call, following Jennifer Nado, the 'Argument from Lack of Other Obvious Options'. Here is an articulation of the argument due to Paul Boghossian:

> (…) we are justified in believing that philosophers appeal to intuitions because they must be - there is *no other viable explanation* of their philosophical practice. Since they take themselves to be justified in making certain sorts of judgments on the basis of thought experiments, and since they are in a good position to see that there is *nothing else* to justify them in making such judgments, a charitable construal of their practice- in line, of course, with all their almost obsessive talk of intuition- would have them appealing to intuitions. (Boghossian 2014, 381; my emphasis)

This essay has three Sections. In Section 1, I argue that the move of invoking intuition as the evidential source in situations where we lack other obvious options amounts to a fallacy. In Section 2, I outline a view according to which arguments serves

[3] As the Argument from Lack of Other Obvious Options is put forward as one of the most plausible arguments for the intuition-centered answer, a rejection of the argument will, of course, be an effective way of undercutting support for the intuition-centered account of thought experiments. For arguments to the effect that intuitions do not play an evidential role in philosophy, see Cappelen (2012), Deutsch (2015) and Machery (2017).

the role of evidence for thought experiment beliefs. In Section 3, I ask whether the considering of a thought experiment in isolation from the text in which it occurs (and so in isolation from arguments) can give rise to additional justification. I argue that it can, and offer three intuition-free accounts of how thought experimentation, in isolation, can give rise to justification.

According to the first account, thought experiment beliefs are justified in virtue of being based on memory. According to the second account, thought experimentation gives rise to new knowledge by means of providing a framework for systematizing tacit beliefs. According to the third account, our thought experiment beliefs are justified in virtue of being formed within a reliable process of imagining.[4]

1. The Argument from Lack of Other Obvious Options

In hope of creating some continuity for the reader, I will center this essay around the so-called 'Truetemp case', a thought experiment coined by Keith Lehrer (in: 1990).[5] Lehrer introduces

[4] This first account is inspired by an interpretation of Thomas Kuhn's theory of thought experiments due to Cooper (2005). The next account draws primarily on what I take to be Kuhn's actual account of thought experiments outlined in Kuhn (1964), but can also be found in Gendler (2010). The last theory is advocated by Gendler (2010) and Williamson (2007, 2015, 2018).

[5] Lehrer's thought experiment is a particularly good case for discussion for at least three reasons. First, it is an example frequently described (by both intuition-friendly and intuition-hostile philosophers) as a paradigmatic example of a philosopher appealing to intuition as evidence. Second, the text in which the thought experiment in question occurs is a work on epistemology; a subfield of philosophy commonly assumed to be particularly intuition-deploying. The third reason is mainly practical. As will become clear below, Lehrer's thought experiment often features in contemporary metaphilosophical debate. See for instance, Sosa (2007), Swain et alt. (2008), Cappelen (2012),

the thought experiment in the context of arguing against Reliabilism, a view according to which reliably produced true beliefs constitute knowledge. The thought experiment goes like this:

> Suppose a person, whom we shall name Mr. Truetemp, undergoes brain surgery by an experimental surgeon who invents a small device which is both a very accurate thermometer and a computational device capable of generating thoughts. The device, call it a tempucomp, is implanted in Truetemp's head so that the very tip of the device, no larger than the head of a pin, sits unnoticed on his scalp and acts as a sensor to transmit information about the temperature to the computational system of his brain. The device, in turn, sends a message to his brain causing him to think of the temperature recorded by the external sensor. Assume that the tempucomp is very reliable, and so his thoughts are correct temperature thoughts. All told, this is a reliable belief-forming process. Now imagine, finally, that he has no idea that the tempucomp has been inserted in his brain, is only slightly puzzled about why he thinks so obsessively about the temperature, but never checks a thermometer to determine whether these thoughts about the temperature are correct. He accepts them unreflectively, another effect of the tempucomp. Thus, he thinks and accepts that the temperature is 104 degrees. Does he know that it is? (Lehrer 1990, 163 f.)

Lehrer, goes on to answer the question negatively. Mr. Truetemp does not know.

The judgment that Mr. Truetemp does not know is standardly taken to carry some rather significant implications. According to Reliabilist theories of knowledge, we ought to attribute

---

Deutsch (2015), Boghossian (2014), Weinberg (2014) and Nado (2016). Familiarity with Lehrer's thought experiment will be advantageous once we encounter those debates.

knowledge to persons that truly believe on the basis of a reliable belief-forming process. Mr. Truetemp's true temperature-beliefs *are* formed on the basis of a reliable belief-forming process. Reliabilism hence predicts the Truetemp case to be an instance that we would classify as knowledge. We do not, however, classify the Truetemp case as an instance of knowledge. Accordingly, Reliabilism appears to have been proven false.

At this point, our epistemological puzzle arises. Our situation appears to be this: Without the input of new empirical information, Lehrer's thought experiment has generated a belief in us that is both new and justified. This new belief appears to have an astonishingly effective evidential force. Confronted with this one imaginary case, one of the most widely endorsed theories of knowledge appears to have collapsed entirely. The questions thus arise: what is the source of our thought experiment belief? What reason do we have to take the thought experiment belief to constitute evidence against Reliabilism?

1.1 Intuition?

Most contemporary analytic philosophers agree that if Lehrer's attempt to refuse Reliabilism is successful it is so because his thought experiment, in a reliable way, generates intuitions that contravene the Reliabilist theory of knowledge.[6] Indeed, the view that intuitions are intended to serve the role as evidence in Lehrer's thought experiment has been assumed to be so obvious that almost no justification has been offered in support of

[6] A growing number of contemporary philosophers believe, or are at least worried that, intuitions are not suited to serve as evidence in philosophy. The worry springs, by and large, from an empirical challenge raised by philosophers publishing under the banner of *negative experimental philosophy*. See Chapter 2 of Machery (2017) for a systematic review of the empirical findings of experimental philosophy.

it. In recent years, however, the view that intuition serves the role of evidence in the Truetemp case and in other thought experiments has been challenged.[7] The Argument from Lack of Other Obvious Options has been offered as a response to these challenges.

The Argument from Lack of Other Obvious Options makes two claims. First, it claims thought experiment beliefs (such as the belief that Mr. Truetemp does not know) to be justified in virtue of being intuitive. Second, it claims intuitions to be the only obvious candidate for serving the role of evidence for thought experiment beliefs. In this essay, my primarily focus will be on the latter claim. I will not, therefore, say much about what intuitions are and what reasons we have for taking them to be obvious candidates for serving the role of evidence for thought experiments beliefs. In order to outline an alternative to the intuition-centered account, as I will do in Section 2 and 3, the topic of intuition should not, however, be avoided entirely. This articulation of intuitive justification, due to David Chalmers, is one I intend to contrast the postulations of this essay with:

> For economy of expression, let us abuse language by saying that a justification is broadly inferential if it is inferential, perceptual, introspective, memorial, or testimonial (…). We can say that intuitive claims have a broadly noninferential justification: justification that does not derive from any of these sources. (Chalmers 2014, 537)

To this minimalistic definition of intuitive justification, we can add a second feature. Intuition-friendly philosophers often claim that "the role and corresponding epistemic status of philosophical intuitions are similar to the role and corresponding epistemic status of perceptions." (Weinberg and Alexander

[7] See Section 2.

2014, 188) Like perception, intuition is commonly thought to provide "non-inferential, defeasible justificatory foundation." (ibid.) Thus understood, intuitions enjoy a privileged justificatory status. They are justifiers that do not stand in need of independent justification. Hence, the story goes, "intuition (and intuition statements) are to thought experiments as perceptions (and observation statements) are to experiments." (Rowbottom 2014, 119)

In arguing for an intuition-free account of how thought experiment believes are justified, this is the kind of theory of intuitive justification I will have in mind. That is, I will seek to articulate an account of how thought experiment beliefs are justified that does not claim thought experiment beliefs to be broadly noninferential, and which does not claim thought experiment beliefs to enjoy some sort of privileged justificatory status.

1.2 An Analogy

Nado presents the Argument from Lack of Other Obvious Options in the following way:

> When I consider the Gettier case, I surely don't have a visual experience upon which I base my subsequent belief that the protagonist of the case fails to know. Nor an auditory experience, nor a memory, nor an introspection and so forth. I seem to simply know, though I cannot say how. Invoking intuition as the evidential source at least takes a step towards an explanation of how this might be so. (Nado 2016, 793)

Nado's line of reasoning appears to apply to the Truetemp case too. Our judgment that Mr. Truetemp does not know does not seem to be based on auditory experience, memory, introspection and so forth. Something different seems to be going on. We might not be capable of explaining *why* we know that Mr. Truetemp does not know, but we are nevertheless confident that

he does not. Should we agree then, with Nado and Boghossian, that we ought to invoke intuition as the evidential source in lack of other obvious options? Would that be a step towards an explanation of how we know that Mr. Truetemp does not know?

I think the answer is no. With no positive evidence to the effect that intuition serves as evidence in thought experiments, it is unclear how invoking intuition would help. As Nado admits (in a footnote), invoking intuition is "not necessarily a very large step of course." (ibid.) It seems to me, however, that invoking intuition as the evidential source *should be* a large step in order for us to be justified in claiming that it answers the Epistemological Puzzle of Thought Experimentation.[8]

Consider the following analogy: Mrs. Garden is found murdered. Mr. Garden is a weird-acting guy with a dubious reputation; friends and neighbors of the Gardens depict him as having a wife-killing vibe. Moreover, Mr. and Mrs. Garden mainly kept to themselves. This makes Mr. Garden the primary suspect in the murder case. There are no other obvious suspects. The police thus conclude that Mr. Garden killed his wife.

The police appear to be committing a fallacy here. The fact that there are no other obvious suspects is not evidence to the effect that Mr. Garden murdered his wife. Mrs. Garden could have had enemies, she could have been the victim of an accident, or she could have been the victim of a serial killer. These may not

[8] The 'Argument from Lack of Other Obvious Options' is put forward as one of the more plausible arguments for an intuition-centered account of thought experiments. The Argument is explicitly endorsed Nado (2016), Boghossian (2014), Chalmers (2014) and Weinberg (2014) Furthermore, it is claimed to be at least tacitly endorsed by most intuition-theorists (Nado 2016, 793.).

be obvious options, but the police would nevertheless be wrong to exclude them. A lack of other obvious suspects should not result in the police sending Mr. Garden to jail without further evidence. Convicting the husband of murder might well be a step *away* from a solution to the murder-case; not a step towards it.

The analogy makes clear, I hope, why we should reject the Argument from Lack of Other Obvious Options. Being short of obvious options is not to be short of possible options. Non-obvious sources of evidence could be at play in thought experiments. Just as other candidates ought to be considered in the murder-case, more alternatives ought to be investigated before a conclusion regarding the Epistemological Puzzle of Thought Experimentation is to be drawn.

1.3 Intuition as an Obviousness Option

To say that there may be other options is not, however, to deny that intuition is an *obvious* option. The fact that Mr. Garden was the only person Mrs. Garden was known to interact with, combined with his weird behavior and his somewhat dubious reputation does make him an obvious suspect in our murder case. In the case of intuition and thought experiments, there are at least two reasons for believing intuitions to be obvious candidates for evidential source.

The first reason is pointed out by Boghossian. Philosophers' "almost obsessive talk of intuition" gives us reason, he claims, to believe that philosophers are appealing to intuition as evidence. (Boghossian 2014, 381) The 'Argument from Intuition-talk' has been thoroughly discussed in the literature, and thus I

won't say much about it here.[9] My view is that intuition-talk in philosophy should not matter much more than town gossip in murder cases. We should of course care about the opinions of those who knew Mr. and Mrs. Garden. If most people will have it that Mr. Garden murdered his wife, that gives us some reason to suspect Mr. Garden. Most people believing Mr. Garden to be a wife-killer is not, however, sufficient evidence for his guilt. The same goes for intuition-talk. We want our answer to the Epistemological Puzzle of Thought Experimentation to be based on more than self-descriptions found in philosophical practice.

The second reasons why intuitions are obvious options is the following: Thought experiment beliefs are, by many, assumed to be accompanied by *a feeling of a special kind*. As put by Gendler, most seem to agree that thought experiment beliefs "does not feel like inference from known premises to inductively or deductively implied conclusions." (Gendler 2010, 43) Moreover, the considering of a thought experiment appears to feel different from testimony. Reaching a judgment regarding Thomson's violin case, for instance, feels different from being told that the president of Guatemala is against abortion. The same appears true for perception. Considering Searle's Chinese room case feels different from looking at a map to see where in China the city Wuhan is located. One reason why we suspect intuition, and not argumentation, perception or testimony, might be the following. Intuitions are assumed to be characterized by a special phenomenology. Platinga, for instance, describes intuitions to have "that peculiar form of phenomenology with which we are all well acquainted, but which I can't describe in any way other than as the phenomenology that goes

[9] See, for instance, Cappelen (2012), Chudnoff (2013) and Weatherson (2014).

with seeing that such a proposition is true.(Plantinga 1993, 105 f.) The same special feeling is described by George Bealer when he writes that a case, when it is first considered, often seems neither true nor false. "After a moment's reflection, however, something happens: it now seems true; you suddenly "just see" that it is true." (Bealer 1996, 5) This is, I think, what Nado is getting at too when she writes that she seems "to simply know" that the protagonist in the Gettier case does not know, "thought she cannot say how." (Nado 2016, 793) On this interpretation, then, intuition presents itself as an obvious option because it appears to be the only option capable of accounting for the characteristic phenomenology of thought experiment beliefs.

In response to this, at least two routes are available. First, we can deny, with Cappelen and Williamson, that thought experiment beliefs come with a special phenomenology. That would be to deny that intuition is an obvious option. Second, we can grant (if only for sake of argument) that their special phenomenology does make intuition an obvious option but deny that this is relevant. In our analogy, the two options amount to something like the following: deny that Mr. Garden has a weird appearance, or deny that Mr. Garden's appearance is of relevancy to whether or not he killed his wife. The latter option seems more fruitful in our murder case. Looking for people with personalities that fits the bill of a murderer is something we stopped doing a long time ago. Sometimes people who appear perfectly normal kill their neighbor. If we limit our suspects to only those that have the personality of a murderer (whatever that means), we might not catch our killer. We don't send a husband to jail for seeming dubious or because the word on the street is that he killed his wife. Better evidence is required. Likewise, we should not draw conclusions without fur-

ther evidence that intuitions serve as evidence in thought experiments. If we require that the correct reply to the Epistemological Puzzle of Thought Experimentation must account for the so-called special phenomenology of thought experiment beliefs, we may not find the right answer. In Section 2, I propose that it is arguments, not intuition that serve as evidence for thought experiment beliefs. Some will reject that proposal on the basis of the alleged special phenomenology being left unexplained.[10] That is, in my opinion, to commit a mistake. It would be, given that my analogy holds, similar to rejecting potential suspects on the basis of them not having the personalities required for murder.

My argument in this essay does not, however, depend on the plausibility of the murder-analogy. Regardless of whether there really *is* something phenomenologically special about thought experiment beliefs, it seems clear that there would be an advantage, at least dialectically, if our solution to the Epistemological Puzzle of Thought Experimentation can account for the characteristic phenomenology claimed to be accompanied by thought experiment beliefs. In Section 3, I outline three theories of the epistemology of thought experimentation that, in my opinion at least, give an account of the special feeling allegedly characteristic of thought experiment beliefs.[11]

[10] Gendler, for instance, rejects a view due to John Norton on the basis that it "requires that something that feels like contemplation of an imaginary scenario is actually the execution of an argument." (Gendler 2010, 43)

[11] With no clear account of what the phenomenology of thought experiment beliefs amount to, it will admittedly be difficult to judge whether these three theories should count as obvious options or not. I return to the issue in Section 3.

2. Argument, not Intuition

In this section I defend a view according to which philosophers doing thought experiments trade in arguments, not intuitions:

> Analytic philosophy is chock-full of hypothetical examples and thought experiments, of course, but analytic philosophers argue for their claims about what is or is not true in these cases and thought experiments. It is these arguments, not intuitions, that are, and should be, treated as evidence for the claims. (Deutsch 2015, XV)

The view that arguments serve the role of evidence for thought experiment beliefs is most notably argued for by Herman Cappelen in *Philosophy Without Intuitions* and by Max Deutsch in *The Myth of the Intuitive*.[12] The two books differ in emphasis, but the overall argumentative strategy is fairly similar. The claim that intuition plays a central evidential role in thought experiments, Cappelen and Deutsch point out, is a straightforwardly empirical claim about how philosophers do thought experiments. Accordingly, the claim can be supported or rejected on the basis of looking at philosophical texts. That is the strategy pursued by Cappelen and Deutsch. Cappelen examines eleven cases assumed to be paradigmatic instances of philosophers appealing to intuitions as evidence.[13] Deutsch, in his book, examines seven additional ones. Neither Cappelen nor Deutsch find textual support for the claim that intuitions are

[12] See Cappelen (2012) and Deutsch (2015). The view that intuitions does not serve as evidence in philosophy is also endorsed by Earlenbaugh and Molyneux (2010), Dorr (2010), Molyneux (2014), Ichikawa (2013, 2016) and Machery (2017).

[13] There are ten case studies in (Cappelen 2012) and one in (Cappelen 2014).

treated as evidence in these texts.[14] What they do find, however, is textual evidence to the effect that philosophers offer arguments for their thought experiment judgments. With no textual evidence to the effect that intuition plays evidential roles in thought experiments, and with eighteen examples of arguments serving the role of evidence, we seem to have ample reason for taking the answer to the Epistemological Puzzle of Thought Experiments to be intuition-free.

This section is divided into two parts. In Section 2.1, I ask whether it is true that Lehrer is making an argument. I argue that he is. In Section 2.2, I reply to one of the more frequently raised objections against Cappelen and Deutsch's view: 'The Relocation Problem'. This is the objection that the argument-not-intuition view is not able to eliminate intuitions from thought experiments. Instead, methodological views such as Cappelen's and Deutsch's merely relocate the place in which intuition is appealed to as evidence.

2.1 Is Lehrer making an argument?

According to Cappelen (2012), arguments are what play the justificatory role in Lehrer's rejection of Reliabilism. Cappelen suggest the following argument to be the *most central* argument for the claim that Truetemp does not know:

> P1: More than possession of correct information is required for knowledge. One must have some way of knowing that the information is correct.

[14] The cases examined by Cappelen and Deutsch are thought experiments typically assumed to be paradigmatic examples of appeals to intuition as evidence. For instance: the trolley problem, the Gettier case, Thomson's violin case, Chalmers' Zombie argument and Kripke' twin earth.

P2: Truetemp has no way of knowing that the information is correct
C: Truetemp does not know

I think Cappelen is right. Critics are, however, unperturbed. To read Lehrer as adducing an argument in favor of his thought experiment judgment is, they maintain, an implausible reading.[15] According to several intuition-theorists, Lehrer's thought experiment is "better understood as Lehrer helping draw our attention to what he takes to be aspects of the case that will produce the relevant cognitive response in us." (Weinberg 2014, 552) The case "helps us have the intuition Lehrer has." (Chudnoff 2017, 383)

Why would Lehrer *not* make an argument, but aim to produce intuitions in his readers, instead? On the intuition-centered account of thought experiments, a theory is plausible to the extent that it is able to account for 'our' intuitions about cases that are relevant to the topic in question. Thus, what serves as evidence for or against a theory of knowledge is what most or all people intuit when confronted with knowledge-related thought experiments. Hence, if 'our' intuition indeed is that Mr. Truetemp does not know, then we do (on the intuition-centered account) have evidence to the effect that Reliabilism is false. That explains Lehrer's alleged interest in producing intuitions in his readers.

This interpretation is claimed by intuition-theorists to be more plausible, indeed even more charitable, than the interpretation according to which Lehrer is making an argument. One reason why we should not read Lehrer as marshalling an argument,

[15] This objection is leveled Weinberg (2014), Boghossian (2014) and Chudnoff (2017).

intuition-theorists claim, is because the alleged argument would be a bad argument. That should, the critics claim, make us suspicious. As put by Chudnoff: "If we find ourselves attributing it to him, charity requires us to step back and ask what else might be going on." (ibid., 381)

Moreover, whereas the intuition-theorist is able to account for why the thought experiment appears in the text, the intuition-denier seems to lack such an explanation. Boghossian puts the objection in the following way:

> (…) it would make an absurdity of the whole point of constructing the thought experiment to think of Lehrer as arguing for the claim that Truetemp doesn't know, by helping himself to the principle that knowledge requires more than correct information. If he already thought of himself as knowing the principle, why would he need to construct an elaborate sci-fi example? (Boghossian 2014, 377)

The intuition-free interpretation is thus faced with two obstacles: the bad-argument objection and the issue of explaining what function the Truetemp case is intended to have, if not the function of generating an intuition. What can be said in response?

The correct response to the bad-argument objection is, in my opinion, to follow Deutsch in denying that the argument actually *is* a bad argument. As Deutsch points out, the argument given for the Truetemp-judgment is not meant to stand on its own. The argument is accompanied by an argument from analogy and this, Deutsch claims, strengthens Lehrer's case. To this observation it should be added (and I return to this point below) that the principle that knowledge requires more than correct information is argued for in several of the chapters of Lehrer's book. Accordingly, Boghossian's complaint misses its target.

Lehrer *reasonably* thinks of himself as knowing the principle because he spends several chapters arguing for it.

One need not, however, agree that Lehrer provides convincing arguments in order to dismantle the bad-argument objection. Another avenue of response open to the intuition-denier is to question the relevancy of the bad-argument objection. For if premises are articulated in the text, and conclusions are drawn from these premises, then there is an argument in the text. Regardless of the quality of that argument, we thus have textual evidence to the effect that it is argument, not intuition that is supposed to do the justificatory work. Whether the arguments do that well or not seems to be beside the point.

What about the second obstacle? The most plausible answer is, I think, that Lehrer thought of the Truetemp case as fulfilling an illustrative function. Lehrer explicitly describes his thought experiment (and other thought experiments appealed to throughout his book) as being illustrative. Here is Lehrer: "A person totally ignorant of the reliability of the process producing his belief would not know that what he believes is true, even if he had no information that would undermine his belief. *The example of Mr. Truetemp illustrates this perfectly*" (Lehrer 1990, 165; my emphasis)

Hence, intuition deniers do have an answer to the question of why Lehrer constructs an elaborate sci-fi example. (Boghossian 2014, 377) As a result, we are left with two interpretations. How do we determine whether Lehrer's thought experiment is illustrating a principle already argued for or whether he is instead constructing the thought experiment in order to trigger relevant intuitions in readers?

An obvious possibility suggest itself: just ask Lehrer what role he intended his thought experiment to play. If we treat Lehrer as the authority on the question, we have ample reason to abandon the intuition-centered reading. First, at several instances, Lehrer describes himself as being in the business of assessing arguments. For instance, immediately after presenting what Cappelen assumes to be the most central argument in the text, Lehrer reminds the reader that "this line of argumentation we have already encountered, in earlier chapters." (Lehrer 1990, 164) Second, immediately after presenting the Truetemp case, Lehrer explicitly says that the example is not meant as a decisive objection and that it should not be taken as such either. (ibid.) Lehrer then goes on to claim that the force of his principle (that is, the principle that serves as a premise in Lehrer's most central argument on Cappelen's reading) does not depend on whether the Truetemp case constitutes a decisive objection or not. *That* statement is in direct conflict with the intuition-centered account. For according to the intuition-theorist, Lehrer's rejection of Reliabilism rests *solely* on intuitions about the thought experiment. Lehrer claims, however, that the fundamental issue remains regardless of the force of the thought experiment. Finally, Cappelen confirms that in conversation, Lehrer denies that he is relying on an intuition and confirms that he thought of himself as making an argument.[16] In conclusion: if Lehrer's thought experiment has the function of generating intuitions, then this fact is and has been opaque to Lehrer himself.

One may object, however, that philosophers are not always the best authority on the question of what, as a matter of fact, they are doing when they do philosophy. Even if we discard these reasons on the basis of the possibility of Lehrer being method-

[16] Personal correspondence.

ologically confused, however, there is reason to prefer the interpretation according to which Lehrer is making an argument. For even if Lehrer believed his thought experiment to generate the relevant intuition, it would be strange for him to stop there. Why wouldn't he also adduce arguments in favor of his judgment? Philosophers usually attempt for their conclusions to be copiously supported. To offer just *one* consideration (that the judgment is either intuitive or not) is not to offer strong evidence. We should expect more of a clearheaded philosopher such as Lehrer, and the fact that the intuition-based reading of the Truetemp case makes Lehrer's rejection come out this flimsy should, I think, make us skeptical of the intuition-based reading. Moreover, what is the difference between an author helping the reader to a conclusion by means of pointing the reader to relevant aspects of a case and an author helping the reader to a conclusion by means of spelling out some premises? It is unclear what the difference between the articulation of relevant aspects and the articulation of premises in an argument is supposed to be.

In any case, I find premises and I find conclusions drawn on the basis of premises in Lehrer's text, as do Cappelen and Deutsch. Perhaps we claim to identify something that is not really there, but for reasons just outlined I find it reasonable to assume that there are arguments in Lehrer's text. In what follows I take a closer look at one of those arguments and I ask what justification there is for holding the premises in that argument to be true.

Before doing so, however, one possible source of confusion should be addressed. Critics have claimed the no-intuition view

to presupposes a false conflict between arguments and intuitions.[17] In doing so, they have attributed an interpretive principle to Cappelen according to which any argument presented for a proposition excludes the possibility of the proposition being based on intuition. (Chalmers 2014, 539 f.) This principle is false, critics claim, for interactions between argument and intuition often is "friendly, even complementary; it is certainly not competitive or mutually excluding." (Bengson 2014, 572)

Here is Cappelen's reply:

> I tried hard, but obviously not hard enough, to make sure that readers wouldn't think I assumed that if an argument is given for p, then p is not supported by an intuition. My view was this: we need evidence that the intuiting of p is doing work. None of my respondents provides such evidence. (Cappelen 2014, 599)

I agree with Cappelen. The reason why we shouldn't assume intuitions to be doing justificatory work is that there is no evidence that they do. If such evidence were provided, however, none of the intuition-deniers (I suspect) would quibble with the claim that intuition and argument interact in friendly and complementary ways. But such evidence has not been provided. Accordingly, we do not have reason to believe that intuition plays an evidential role in thought experiments. Neither do they need to, at least not in the case discussed in this essay. Lehrer's rejection of Reliabilism is compelling without appeal to intuition.

2.2 The Relocation Problem

[17] The objection is most notably made by Bengson (2014) and Chalmers (2014).

According to the Relocation Problem, evidence to the effect that it is arguments not intuitions that serve as evidence in philosophical texts is not enough to conclude that intuitions do not play an evidential role in thought experiments. What Cappelen and Deutsch have illustrated is, at most, that there is no appeal to intuition as evidence at *a particular level.* Even if one grants that it is not intuition but arguments that justify judgments about thought experiments, one may still be worried about the premises in those arguments. How do we know that the premises in arguments for judgments about thought experiments are true?

One tempting response, suggested by intuition-friendly and intuition-hostile philosophers alike, is that "in many cases the premises in philosophical arguments are based on intuition." (Bengson 2014, 571)[18] Here is an articulation of the Relocation Problem due to Jonathan Ichikawa:

> Cappelen is quick to emphasize that there are arguments underwriting my judgment about Mr Truetemp- but arguments proceed on the basis of premises, and what story are we to tell about my epistemic access to the relevant premises? (...) Insofar as it doesn't seem very plausible that perceptual experience can ultimately be establishing the premises from which I can conclude that Mr Truetemp does not know, one might be tempted to think that it must be some other kind of experience, which plays a similar role to that of perceptual experience. (Ichikawa 2013, 115 f.)

Nado levels the same objection:

[18] This claim is made explicitly by Chalmers (2014), Nado (2016, 2017) and Chudnoff (2017). It is also hinted at by Bengson (2014) and by Climenhaga (2018).

> How does Cappelen's characterization do anything more than push the problem back one step? The principle that Cappelen takes to be a premise is in just as much need of justification as the claim that Mr. Truetemp does not know. How is it to be justified, other than via intuition? (Nado 2016, 796)

What can be said in response to this objection?

A reply to the relocation problem can, in my opinion, be found in the wider context of Lehrer's book. Once we consider Lehrer's argument as a whole, and not merely the few passages in which the Truetemp case appears, we see that *further argumentation* is what establishes the premises of Lehrer's arguments against Reliabilism.

Let's take a closer look at Lehrer's most central argument concerning Mr. Truetemp, as proposed by Cappelen. According to the second premise, *Mr. Truetemp has no way of knowing that the information he receives is true.* Note, that this is merely a feature of the case. Mr. Truetemp is described as having no idea that a device was implanted in his head, he is described as not ever checking whether his temperature thoughts are correct, and his unreflective acceptance of temperature thoughts is described as an effect of the device. What is important for our purposes, however, is that Lehrer nevertheless offers reasons in support of this premise. Lehrer explicitly states that Mr. Truetemp "did not consider any evidence concerning the matter, and *that is why* he does not know that his thoughts about the temperature are correct." (Lehrer 1990, 165; my emphasis)

Lehrer's first premise has been the more contentious one: *More than possession of correct information is required for knowledge. One must have some way of knowing that the information is correct*. As we have seen, both Nado, Boghossian and

Ichikawa claim this principle to be entirely unsupported. They are, to some extent, right. The objection nevertheless misses its mark. This is because the objectors fail to recognize what Lehrer is really trying to do in *A Theory of Knowledge*. Lehrer is not aiming to establish the principle that more than correct information is required for knowledge. This principle is taken for granted throughout his book.

Here's how Lehrer helps himself to the premise. In Chapter 1, Lehrer points out that there are multiple sorts of knowledge. Accordingly, to seek a general analysis of knowledge is to set oneself a too unspecific goal. One ought to, as Lehrer points out, specify what sort of knowledge one's analysis of knowledge is an analysis of, prior to inquiry. Lehrer does so already at page 3 in his book, where he makes clear that he will be concerned with *knowledge in the sense associated with scientific inquiry*.[19] In order to do science well, he maintains, it is not enough to merely have correct information. To make progress in science, Lehrer writes, "one must be able to tell whether one has received correct information or not." (ibid., 4)

According to the objectors, Lehrer's principle must be rooted in intuition because it cannot be the case that Lehrer just "help himself" to the principle. But Lehrer does "help himself" to the principle; he takes it for granted that his readers will be interested in the sense of know "that in which 'to know' means to recognize something as information." (ibid., 3) Lehrer need

[19] Lehrer takes the type of knowledge analyzed in his book to be a "more significant kind of knowledge" as compared to other sorts of knowledge. This is primarily due to its practical role. Analyzing the sort of knowledge required for science is, Lehrer assumes, more important than analyzing other sorts of knowledge.

not, then, appeal to intuition as evidence for his thought experiment judgment. The judgment that Mr. Truetemp does not know is justified in virtue of being based on an argument in which the first premise is an argumentative starting point and where the second finds support in the text. This more holistic reading of the thought experiments gives us a clue as to what the function of Lehrer's thought experiment may be. The thought experiment seems to have, at least primarily, an illustrative function. If we accept that more than correct information is required for knowledge (of the type central to science), what follows? What the Truetemp case successfully illustrates is, I think, that Reliabilism is not a theory of knowledge consistent with the scientific sense of knowledge according to which more than correct information is required for knowledge. This result is, if my reading is correct, attained without appeal to intuition as evidence.

3. Thought Experiments in Isolation

The reply outlined thus far emphasizes the importance of considering the context in which thought experiments occur. Once we do, we see that philosopher's claims about thought experiments are argued for. Hence, we have a reply to the epistemological Puzzle of Thought Experimentation. The new knowledge acquired in a thought experiment situation is deductively implied by arguments marshalled by the author of the thought experiment.

To say that thought experiment judgments are justified in virtue of being based on argument is not, however, to say that *readers* form their thought experiment beliefs on the basis of considering an argument. More typically, it seems, we form thought experiment judgments on the basis of considering thought experiments in isolation. When doing so, we typically take ourselves to be justified in holding our thought experiment beliefs to be

true. We do not, it seems, need to read Lehrer's book in order to be justified in taking our belief that Mr. Truetemp does not know to be true.

The view that thought experiments considered in isolation can give rise to additional justification for the content of thought experiment beliefs strikes me as correct. In this section, therefore, I consider three ways in which the considering of a thought experiment, in isolation, can give rise to additional intuition-free justification for thought experiment beliefs.

3.1 Cooper's Kuhn: Thought Experiments as Mnemonics

Here's a simple exercise. Let's say I ask you who your classmates were in 5th grade. How many names do you remember? If you are like me, you won't remember more than a few names. You nevertheless *know* who all of your classmates were. If you run into a person you once went to school with, this person is not going to be a total stranger. You remember the person even though you might not have his name on immediate recall. Now, consider the following questions. Did any of your classmates live in your neighborhood? Did you do sports with any of them? Did you play music with anyone from your class? Mnemonics of this sort might make you imagine your neighborhood, the pitch where you played football growing up, or a school musical. Plausibly, you come to remember more names. The considering of a few relevant questions help you tease out information previously forgotten.

According to Rachel Cooper, Kuhn holds a similar process to be at play when we are doing thought experiments. Thought experiments are not, on this view, tools for acquiring *new* knowledge of the world. Instead, thought experiments function as a sort of mnemonic. Their narrative structure can, in some

cases, trigger the memory of the reader, and thus draw out information that was not previously available. This provides a simple solution to the Epistemological Puzzle of Thought Experimentation. Since the knowledge acquired in a thought experiment is merely retrieved knowledge, the Epistemological Puzzle of Thought Experimentation evaporates. (Cooper 2005, 330 f.)

As will become clear below, I don't think that this view is plausibly attributable to Kuhn. Thought experimentation *does*, on Kuhn's view, lead to new knowledge. I nevertheless like the idea of thought experiments as tools for drawing out forgotten knowledge. It seems plausible that the judgment that Mr. Truetemp does not know could be based on memory.

Here's the mnemonical reading of the Truetemp case: You know a lot of things about knowledge. Some of your knowledge about knowledge you know explicitly. You know, for instance, that knowledge requires a high level of certainty, and you know that some routes to knowledge are more reliable than others. Some of your knowledge about knowledge you have forgotten. Among the things you know about knowledge is, presumably, the principle that more than correct information is required for knowledge.

To see the plausibility of that claim, consider the following anecdote. As a kid, I had a rhyme for remembering that seven multiplied by eight equals fifty-six. When asked about this particular equation I consistently gave the right answer. This was, however, *before* I knew how multiplication worked. I was not, for instance, in a position to figure out what six multiplied by eight was. Accordingly, my parents did not (at least I hope they did not) go around bragging about how I knew how to do multiplication. They knew that there was a distinction between

knowing something and merely being in possession of right information. Indeed, most people do. It is standardly taught in school; when doing mathematics in elementary school, for instance, one is taught that having the correct answer does not count for much unless one can "show the work".

When you are asked to consider the Truetemp case, then, you do not come empty handed. You know, or at least you used to know, that more than correct information is required for knowledge. Accordingly, you do not form the *new* belief when confronted with the Truetemp case. Your thought experiment belief is instead based on memory. Since the knowledge is not, contrary to what the intuition-theorist claims, new, the Epistemological Puzzle of Thought Experimentation does not arise.

This account of Lehrer's thought experiment strikes me as plausible. The view that thought experiments function as mnemonics is, however, rather limited. Thought experiments, on the mnemonical view, can only be successful if the thought experimenter has all the relevant information already stored in her memory. For a lot of thought experiments, however, that is unlikely to be true. What can be said about those thought experiments? A different (but in my opinion more accurate) reading of Kuhn can, I think, provide some answers.

3.2 Thought Experiments as Tools for Detecting Conflict

Contrary to what Cooper claims, Kuhn does not endorse a view according to which thought experiments have the function of helping readers retrieve forgotten knowledge. Quite on the contrary: Kuhn explicitly denies the view that "a thought experiment can teach us nothing that was not known before." (Kuhn 1964b, 252) We do, on Kuhn's view, learn something new when we are doing thought experiments.

There are, Kuhn claims, close similarities between the function of thought experiments and the function of actual laboratory experiments. Concerning the role of thought experiments in science, Kuhn writes:

> Historically their role is very close to the double one played by actually laboratory experiments and observations. First, thought experiments can disclose nature's failure to conform to a previously held set of expectations. In addition, they can suggest particular ways in which both expectation and theory must be henceforth revised. (ibid., 261)

The idea that thought experiments can disclose nature's failure to conform to previous expectations needs some unpacking. First, however, it is worth highlighting that Kuhn's view is merely a view of how *some* thought experiments function. Importantly, Kuhn does not claim all thought experiments to function as tools for detecting and revising mismatches between expectations and nature. His aim is rather to describe a category of thought experiments central to science.[20] As put by Kuhn: "No single thought experiment can, of course, stand for all of those which have been historically significant. The category "thought experiment" is in any case too broad and too vague for epitome." (ibid., 241) I agree with Kuhn that we should not

[20] Note that thought experimentation is not some strange activity that only philosophers indulge in. In fact, appeals to thought experiments occur within most (if not all) intellectual disciplines. Kuhn takes the role of thought experimentation in science to be particularly central. Indeed, Kuhn describes thought experiments as one of the essential analytic tools deployed in crisis science. By bringing about conceptual reform, a thought experiment can, Kuhn claims, trigger scientific revolutions.

seek a unifying account of thought experiments.[21] I nevertheless think Kuhn's account of thought experiments as tools for detecting error could explain a wider category of thought experiments than suggested by Kuhn himself. In particular, I believe Kuhn's account can be extended to some philosophical thought experiments.

The claim, then, is that *some* thought experiments function as tools for disclosing nature's failure to conform to a previous set of expectations. What does Kuhn mean by this? Take a simple example. Sam has never seen nor heard of black swans before. Accordingly, he believes swans to be white. As Sam sees a black swan for the first time, then, what Sam experiences is a mismatch between his expectations of what swans are like and what swans are really like. Nature fails to conform to his expectations that swans are white.

Kuhn claims thought experiments to function in a similar way. The idea, simply put, is that the person considering a thought experiment does not enter the situation neutrally. We enter the thought experiment situation with a set of beliefs about what the world is like. In cases where our beliefs are false or inaccurate, a thought experiment can help us discover the mismatch between our expectations of the world and what the world is really like.

[21] Cooper rejects Kuhn's account on the basis that it cannot be applied to *all* thought experiments. Simplicity, Cooper claims, "dictates that a common account of all thought experiments should be sought if at all possible." I don't think that a unifying account of thought experiments is possible or worth aspiring to. I thus disagree with Cooper. The fact that Kuhn's theory cannot account for all thought experiments is not a reason to reject it.

There is, however, an important difference between Sam discovering that it's not true that all swans are white and our acquiring of new knowledge in a thought experiment situation. Sam learns that his conviction that all swans are white by means of seeing a black swan. His new knowledge is based on sensory experience. This is not, however, the case for new thought experiment beliefs. There is, as Kuhn stresses, no new sensory input. How, then, can thought experiments play the role of detecting and correcting mismatches between what we think the world is like and what it is really like? Kuhn's reply is the following:

> Laboratory experiments play these roles because they supply the scientist with new and unexpected information. Thought experiments, on the contrary, must rest entirely on information already at hand. If the two can have such similar roles, that must be because, on occasion, thought experiments give the scientist access to information which is simultaneously at hand and yet somehow inaccessible to him." (ibid., 261)

At this point, it is worth highlighting the main difference between my reading and Cooper's reading of Kuhn. According to Cooper, the information which is 'simultaneously at hand and yet somehow inaccessible' is forgotten knowledge. I think this is inaccurate. Whereas forgotten knowledge may play a role in the acquiring of new knowledge in a thought experiment situation, the resulting knowledge is *not* merely forgotten knowledge. Doing thought experiments may, if done successfully, lead to knowledge that we did not have prior to considering the thought experiment.

To see how one may come to learn something new without the input of new information, consider the following scene. A few years ago, I played college soccer in Georgia. Most of my

coaches and teammates were Europeans; a total of eight European countries were represented. One of my American teammates thought, however, for a surprisingly long time, that Europe was a country and accordingly that we all came from the same country. A very simple question made her realize the absurdity of that belief: 'why would we all speak English to each other then?' Knowing that people from the same country usually share a common language, my teammate quickly retreated her claim. She was also quick to admit other reasons why she should have known, such as seeing different European flags and being introduced to different types of European food. She had sufficient tacit and explicit knowledge to be in a position to know that her statement was false. All she needed was to reflect on it.

That Europe is not a country was not knowledge my teammate had simply forgotten. She had not, prior to this occasion, held that belief. She was nevertheless *in a position to know* it. Her coming to know that Europa is not a country did not, then, require new empirical data. It merely required an act of reflection.

Let's consider whether the Kuhnian story can be applied to the Truetemp case: Prior to considering the Truetemp case, you have a set of tacit and explicit ideas about what knowledge is. Some of those ideas may be in conflict with each other. For instance, from your day-to-day experience you may know (tacitly or explicitly) that more than correct information is required for knowledge. You may also, however, hold the view (tacitly or explicitly) that reliably true beliefs suffice for knowledge. Having these two ideas baked into your conception of knowledge you are, prior to considering the Truetemp case, liable to be confused. Once you encounter the Truetemp case, the feeling of paradox hits you. You realize that your conviction that reliability produces true beliefs suffices for knowledge

is in conflict with your judgment that Mr. Truetemp does not know. One belief, then, must be jettisoned. Either Mr. Truetemp does know, or the Reliabilist concept of knowledge is false.

3.3 Knowledge by Imagination

The theories outlined thus far all agree that thought experiment beliefs are different from laboratory experiment beliefs in that the latter is based on new empirical input, whereas the former is not. The theory outlined in this section denies that thought experiments differ from laboratory experiments in this respect. By performing a thought experiment, we do get access to new information. New information arises within the process of imagining, and it is on the basis of this new imagination-based information that our thought experiment beliefs arise.

The idea of knowing by imagination might, at least on first glance, seem dubious. For, as Williamson points out: "Imagining is often contrasted with knowledge. When you know nothing about something, you have to imagine it instead. Knowledge deals in facts, imagination in fictions. (cf. Williamson 2015, 113) This common conception of the imagination is, however, inadequate. Here are some mundane examples of how we know counterfactuals via imagination: I could have listened to Mozart while writing all the words of this essay. I could have written the essay in Brazil. The essay could have been one page shorter. I could have written the essay only using my left pinky finger. I have never tried to do these things, but I nevertheless know that I am capable. The central point is this: one does not always have to make something actual to know that it is possible.

Influential versions of the view that imagining can lead to knowledge are defended by Williamson and Gendler. The two

advocate slightly different views of how the imagination can be a source of knowledge, but their views can nevertheless be said to be broadly similar. They agree, for instance, that one central way in which the imagination gives rise to new knowledge is by manipulation of mental imagery. To see how contemplation and manipulation of a mental image can give rise to new knowledge, let's consider a question posed by Gendler: If you removed all the furniture of your next-door neighbor's living room, could four elephants fit comfortably inside? (cf. Gendler 2010, 46) You answer this question, Gendler proposes, by first forming a proportionately sized mental picture of your neighbor's living room. Next, you form proportionally sized images of four elephants. You then block the space that the elephants would occupy. Having manipulated this mental picture, you are now able to *see*, so to say, whether the elephants fit in your neighbor's living room or not. On the basis of this mental image, you form a belief concerning the possibility of four elephants fitting into your neighbor's living room.[22]

A second way in which the imagination can give rise to new knowledge is by means of mental stimulation.[23] By acts of imagination, we seem to be able to put ourselves in other people's shoes. In doing so, we seem to be able to stimulate other people's mental processes. On the basis of this ability we are, to a large extend, able to think and decide on the basis of their beliefs and desires. This latter ability seems to be at play when we consider the Truetemp case. In considering the Truetemp case, you presumably put yourself in Mr. Truetemp's shoes. You imagine what it would be like to be in the situation outlined by the

[22] Williamson offers similar examples. See for instance, Williamson (2007, 142).

[23] For more on mental stimulation and knowledge by imagination, see Kind and Kung (2016).

thought experiment, and from that perspective, you ask yourself whether you know that the temperature is 104 degrees. Presumably, you do not believe yourself to have knowledge it the case you have been invited to imagine yourself in. Accordingly, you judge the Truetemp case to be a case in which it would be wrong to attribute knowledge.

On the knowledge-by-imagination view of thought experiments, we are licensed to take our judgments made within the process of imagining to be true on the basis of the process being a reliable belief-forming process. One reason why we should take the process to be reliable, both Williamson and Gendler points out, is the fact that our imaginations "can in principle exploit all our background knowledge in evaluating counterfactuals." (Williamson 2007. 143) Gendler puts the point in the following way:

> Framed properly, however, a thought experiment can tap into it, and- much like an ordinary experiment- allow us to make use of information about the world which was, in some sense, there all along, if only we had known how to systematize it into patterns of which we are able to make sense. (Gendler 2010, 39)

The idea is that our experience with the world, our sense of what the world is like and the abilities we have developed on the basis of that experience and knowledge, put us in a position to make reliable judgments within the process of considering imaginary scenarios. When confronted with a knowledge-related thought experiment, Williamson claims, we engage in an act of imagination, and we make a judgment "on the basis of an offline application of our ability to classify people around us as knowing various truths or as ignorant of them, and as having or as lacking other epistemologically relevant properties." (Williamson 2007, 118) The fact that our judgments made within

the process of imagination are informed and influenced by our background knowledge and abilities gives us good reason, Williamson and Gendler claims, to take the judgments to be true.

Furthermore, there are (as pointed out by Williamson) evolutionary reasons for us to have developed good imaginations. A good imagination "alerts us to future possibilities, so we can prepare for them in advance- guard against dangers, be prepared to take advantage of opportunities." (Williamson 2019, 58) Evolutionary pressure has plausibly made our imagination selective in that it does not generate too many possibilities, and reality-oriented in that it only suggests scenarios that are actually likely to happen. In this sense, knowledge by imagination is closely linked to (or is even a type of) inductive knowledge.

3.4 Taking Stock

If what I have argued in this section is correct, then the considering of a thought experiment in isolation from the rest of the text in which it occurs can give rise to beliefs that we are justified in taking to be true. More must be said, however, in order to properly vindicate these accounts of the epistemology of thought experiments. More should, in addition, be said about the connection between the different intuition-free alternatives theories outlined in this section. One view that I find plausible is that different processes are at play in different thought experiments. An alternative view is that different processes are at play in different people considering the same thought experiment. Perhaps the belief concerning Mr. Truetemp is, for one person based on a memory and for another based on a process of imagining? A third, and perhaps more plausible alternative, is that our thought experiment beliefs are based on a combination of the three views outlined in this section.

Before concluding, two obstacles are worth considering. First, can the theories outlined in this section account for the special phenomenology allegedly distinct for thought experiment beliefs? Admittedly, I am not entirely sure what the special phenomenology is supposed to amount to. That being said, experiences of remembering, discovering contradictions, making inferences and imagining do, at least on occasions, come with something like a special feeling. Consider the experience of having something on the tip of your' tongue and then suddenly remembering what you had forgotten. Or consider the phenomenology of understanding; an experience of 'pieces coming together.' Discovering or rediscovering knowledge can, it seems, be said to be accompanied by special feelings.[24]

Second, some may be inclined to object that the accounts outlined in this section are not intuition-free alternatives, but rather three slightly unorthodox theories of how intuitions can play an evidential role in thought experiments. Consider, for instance, the view due to Darrell Rowbottom, according to which "thought experiments (and the intuitions therein) rest, ultimately, on experience", where experience "includes learning how to use words, e.g. ostensive definition." (Rowbottom 2014, 120) Or consider a view due to Nevin Climenhaga, according to which judgments made on the basis of broadly non-inferential sources of evidence count as intuitions in cases

[24] This might, of course, not be what the intuition-theorists have in mind when claiming thought experiment beliefs or intuitions to come with a special phenomenology. If so, that is okay. As pointed out in Section 2, people who lack the appearance of a murder (whatever that means) may nevertheless be murderers. Likewise, sources of evidence not capable of accounting for the special feeling distinctive for thought experiment beliefs (whatever that means) may nevertheless serve as evidence in a thought experiment.

where the agent is not explicitly aware of the fact that her judgment is based on a broadly noninferential source of evidence. (cf. Climenhaga 2018) The three accounts outlined in this section appear to be compatible with these views. Hence, the question arises: why say that the three accounts are intuition-free replies to the Epistemological Puzzle of Thought Experimentation and not, as Rowbottom and Climenhaga claim, intuition-centered accounts of how thought experimentation can give rise to new knowledge?

The reason why we should label the three accounts outlined in this section intuition-free replies to the Epistemological Puzzle of Thought Experimentation is, in my opinion, the fact that they are in direct conflict with what I take to be a more central notion of intuition, according to which intuitive justification is broadly noninferential justification with a privileged justificatory status. To label judgments that are based on broadly inferential sources of evidence such as memory, background knowledge, training, inference and so on, seems, therefore, highly counterproductive.

Conclusion

The aim of this essay has been (or was at least intended to be) modest. I have tried to establish that intuition is not the *only* available answer to the Epistemological Puzzle of Thought Experimentation. If what I have argued in this essay has been correct, then the Argument from Lack of Other Obvious Options is false. Thought experiments can provide new knowledge without appealing to intuitions as evidence.

According to the view outlined in Section 2, claims made about thought experiment are justified in virtue of being deductively implied by arguments made in the text in which the thought experiment occurs. On this account, the function of a thought

experiment is primarily illustrative. If what I have argued in Section 3 is correct, however, thought experiments have an additional evidential role. Thought experiment beliefs can be justified in virtue of being informed and disciplined by experience, perception, memory, background knowledge, training and other intuition-free sources of evidence.

One point I've been trying to make throughout the essay is this: we do not come empty handed when we step into a thought experiment situation. Our thought experiment judgments are informed by what we already know. This takes away some of the mystery surrounding thought experiment beliefs. Thought experiments are not telescopes into the abstract realm or "oracles guiding us or misguiding us from the depth". (Williamson 2018, 61) Thought experiments are (among other things) tools for explicating what we already know, frameworks for organizing our tacit commitments and invitations to engage in imaginative exercises. We form our thought experiment beliefs on the basis of quite ordinary capacities: our capacity for retrieving knowledge, our capacity to detect contradictions and our capacity to imagine hypothetical scenarios.

In conclusion, it is worth emphasizing the fact that doing thought experiments do not always yield knowledge. In some cases, one simply lacks the relevant background knowledge. In other cases, one's reflection or imagination does not lead to a clear answer. The method of thought experimentation is thus, although mostly reliable, a fallible method. However, as put by Williamson:

> This isn't a reason for not using thought experiments, for all human faculties are fallible. Rather it's a reason for spreading our bets, not relying exclusively on thought experiments. If we use other methods too, they may help us catch our occasional

> mistakes in judging thought experiments, even if those mistakes are species-wide. Developing systematic general theories, supported by the evidence, is a good way of doing that. (ibid., 65)

I agree with Williamson. We should not expect thought experiments alone to do the justificatory work in philosophical texts. Fortunately, philosophers very rarely (if at all) do. According to Cappelen's and Deutsch's empirical data, philosophers -at least good philosophers- do not restrict themselves to merely pointing out *one* way in which their theories and claims are supported. Lehrer's book illustrates this point very clearly. Through the course of his book, Lehrer develops a systematic (some would even say overly systematic) theory. In doing so, he appeals to a number of different arguments, examples, thought experiments and methods. It is primarily on the basis of this systematic theory that we are justified in judging Reliabilism to be false. On this understanding, our judgments concerning the Truetemp case have a much more limited role than what is standardly assumed. The thought experiment is not, as Lehrer himself points out, a necessary component of the argument against Reliabilist theories of knowledge. That is not to say, however, that the thought experiment is unimportant. As pointed out in Section 2, the thought experiment plays an important illustrative function in Lehrer's text. And as argued in Section 3, the thought experiment can, in an intuition-free way, provide additional justification for the view that Reliabilism is a problematic theory of knowledge. Invoking the Truetemp case strengthens Lehrer's case against Reliabilism. It does so, if what I have argued is correct, without appeal to intuition as evidence.[25]

[25] This essay has benefited greatly from discussion with and comments from many people. I'm particularly grateful to Jesse Tomalty

and Herman Cappelen for carefully reading through and giving feedback on a previous draft of this contribution. It made a great difference. I also want to thank Bernt Ivar Barkved, Audun Syltevik and Evelyn Erickson for commenting on several drafts, and for the hours spent trying to make sense of metaphilosophical problems over the past year.

Bibliography

Bealer, G. "On the Possibility of Philosophical Knowledge". *Philosophical Perspectives*, 10, 1996, 1-34.

Bengson, J. "How Philosophers Use Intuitions and 'Intuition'". *Philosophical Studies*, 171(3), 2014, 555-576.

Boghossian, P. "Philosophy Without Intuitions? A Reply to Cappelen". *Analytic Philosophy*, 55(4), 2014, 368-381.

Booth, A. R. and Rowbottom, D. P. (eds.). *Intuition*. Oxford: Oxford University Press, 2014.

Brown, J. R. "Peeking into Plato's Heaven". *Philosophy of Science*, 71(5), 2004, 1126-1138.

Capellen, H. *Philosophy Without Intuitions*. Oxford: Oxford University Press, 2012.

Cappelen, H. "Replies to Weatherson, Chalmers, Weinberg and Bengson". *Philosophical Studies*, 171(3), 2014, 577-600.

Chalmers, D. "Intuitions in Philosophy: A Minimal Defense". *Philosophical Studies*, 171(3), 2014, 535-544.

Chudnoff, E. *Intuition*. Oxford: Oxford University Press, 2013.

Chudnoff, E. "The Reality of the Intuitive". *Inquiry*, 60(4), 2017, 371-385.

Climenhaga, N. "Intuitions are Used as Evidence in Philosophy" *Mind*, 127(505), 2018, 69-104.

Cooper, R. "Thought Experiments". *Metaphilosophy*, 36(3), 2005, 328-346.

Dorr, C. "Review of 'Every Thing Must Go. Metaphysics Naturalized' by Ladyman and Ross", https://ndpr.nd.edu/news/every-thing-must-go-metaphysics-naturalized/, 2010, accessed 08/06/2020.

Deutsch, M. *The Myth of the Intuitive*. Cambridge MA: MIT Press, 2015.

Deutsch, M. "Replies to Commentators". *Inquiry* 60(4), 2017, 420-442.

Earlenbaugh, J. and Molyneux, B. “If Intuitions Must be Evidential Then Philosophy Is in Big Trouble”. *Studia Philosophica Estonica*, 2, 2010, 35-53.

Gendler, T. S. *Intuition, Imagination and Philosophical Methodology*. Oxford: Oxford University Press, 2010.

Ichikawa, J. J. “Philosophy Without Intuitions”. *International Journal of Philosophical Studies*, 2(1), 2013, 111-116.

Ichikawa, J. J. “Intuitive Evidence and Experimental Philosophy”. J. Nado (ed.): *Advances in Experimental Philosophy and Philosophical Methodology*. London: Blomsburry, 2016.

Kind, A. and Kung, P. *Knowledge Through Imagination*. Oxford: Oxford University Press, 2016.

Kuhn, T. *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: University of Chicago Press, 1964a.

Kuhn, T. “A Function for Thought Experiments”. T. Kuhn: *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: University of Chicago Press, 1964b, 240-265.

Lehrer, K. *Theory of Knowledge*. London: Routledge, 1990.

Machery, E. *Philosophy Within its Proper Bounds*. Oxford: Oxford University Press, 2017.

Molyneux, B. “New Arguments that Philosophers Don’t Treat Intuitions as Evidence”. *Metaphilosophy*, 45(3), 2014, 441-461.

Nado, J. “The Intuition Deniers”. *Philosophical Studies*, 173(3), 2016, 781-800.

Nado, J. “Demythologizing Intuition”. *Inquiry*, 60(4), 2017, 386-402.

Plantinga, A. *Warrant and Proper Function.* New York: Oxford University Press, 1993.

Rowbottom, D.P. “Intuitions in Science: Thought Experiments as Argument Pumps”. A.R. Booth and D.P. Rowbottom (eds.): *Intuition*. Oxford: Oxford University Press, 2014, 119-134.

Sosa, E. “Experimental Philosophy and Philosophical Intuition”. *Philosophical Studies*, 131(1), 2007, 99-107.

Sosa, E. “Intuitions: Their Nature and Probative Value”. A.R. Booth and D.P. Rowbottom (eds.): *Intuition*. Oxford: Oxford University Press, 2014, 36-49.

Swain, S., Alexander, J. and Weinberg, J.M. “The Instability of Philosophical intuitions: Running Hot and Cold on Truetemp”. *Philosophy and Phenomenological Research*, 76(1), 2008, 138-155.

Weatherson, B. “Centrality and Marginalization”. *Philosophical Studies*, 171(3), 2014, 517-533.

Weinberg, J. “Cappelen Between a Rock and a Hard Place”. *Philosophical Studies*, 171(3), 2014, 545-553.

Willliamson, T. *The Philosophy of Philosophy*. Oxford: Wiley-Blackwell, 2007.

Williamson, T. “Philosophical Criticism of Experimental Philosophy”. J. Sytsma and W. Buckwalter (eds.): *A Companion to Experimental Philosophy*. Hoboken: Wiley Blackwell, 2015, 22-36.

Williamson, T. “Knowing by Imagining”. A. Kind and P. Kung (eds.): *Knowledge Through Imagination*. Oxford: Oxford University Press, 2016, 113-23.

Ragnhild Iveranna Hogstad Johrdal

Constructed Kinds that Reflect Real Categories. On Hacking's Looping Kinds and His Use of Autism as an Example

1. Introduction

How can we understand classifications—such as certain psychiatric diagnoses— that change over time? Ian Hacking speaks of interactive or human *kinds* and presents a framework for understanding and explaining how such classifications may change over time through a phenomenon he calls *looping*. Looping entails that the individuals who are classified in certain ways are somehow changed by being categorised, and that these changes, in turn, will lead to the category *itself* changing. This article focuses particularly on an example used by Ian Hacking in his (1995) and (2007), namely the example of autism, and how the change seen in the autism diagnosis, and in our general knowledge about autism, may be said to be a case of looping.

I will commence, in section 2, by discussing categorisation in terms of *kinds*, and by introducing the phenomenon of *looping*. In order to discuss Hacking's autism example, I present a short introduction to the history of the autism diagnosis in section 3, before moving to a framework of looping and how this may be understood in the autism case in section 4. In section 4.1, I pre-

sent the case of the *high-functioning autistic person* as a particular *kind of person*, and in section 4.2, I argue that Hacking's exclusion of other autistic people in this example is not well-founded.[26]

I end, in section 5, with some remarks about how looping can contribute *both* to the creation of new kinds, but also to the destruction of already existing kinds. I also point to areas of research where there is need for further work, also from philosophers.

2. Classification in terms of Kinds

When we speak of *kinds* we should be aware that there are several different interpretations of the concept available. Some philosophers, like (Ellis 2002), speak of strict *natural kinds*, and restrict them to those kinds we find in chemistry and physics—things like electrons and oxygen. Other philosophers broaden this definition somewhat, and say that a kind is 'natural' if it corresponds to a grouping which reflects the structure

[26] I use the phrase 'autistic person' instead of saying 'person with autism'. The reason for my choice of using *identity first* language, rather than *person first*, is, most importantly, that this is the way *most* autistic people prefer to be addressed (see e.g. Kenny et al. 2015). In addition, the choice is related to how the phrase 'person with autism' may be interpreted; it may give the impression that autism is an add-on, something extra in addition to the person, or even something that can be removed.
I will also be using some 'functioning level' language, even though I do not myself support the use of such language. One of Hacking's key examples is that of the 'high-functioning autist', so I will use that terminology here, simply because it is hard to describe his example without using this phrase.

of the natural world.[27]These definitions are too narrow for the questions I will consider here. For the moment, I will simply think of kinds in the following way: as properties that cluster in the sense that some subset of properties are good indicators of the presence of the others.[28] This is quite a vague definition, but I do not want to say too much about *kindhood* as such. For the purpose of this article, a vague idea of what classification in terms of *kinds* amounts to is sufficient. The categorisations I shall discuss are those found *within* the human species.

This means that the relevant categories are those we in general find in the human sciences, including the social sciences, psychology, and psychiatry.

I hold that the specific kinds I will discuss in this article *do* describe real divisions in nature, even though they may not be completely discrete, and *even though* the kinds might *in some sense* be both constructed by us *and* highly flexible. It is this flexibility which points towards the main topic of the article, namely, kinds which are changeable over time. These kinds may be referred to as *interactive kinds*. Even more specifically, I will focus on the subset of interactive kinds which are relevant to humans. A particularly interesting phenomenon occurring among these human kinds is what Ian Hacking refers to as *looping*.

When looping occurs, the people who are classified will influence the category they are in to such a degree that the *category itself* may change. Hacking has referred to this as a form of dy-

[27] Such a definition of natural kinds is found e.g. in (Bird and Tobin, 2018).

[28] We find similar ideas in Bird (2009, 6).

namic nominalism, and he argues that the phenomenon of looping is something which distinguishes the human kinds from other kinds—it is a mark of the human kinds (cf. 1999, 108). This has been challenged by, for example, Muhammad Ali Khalidi, who states that interactive kinds may be found outside the scope of the human kinds (cf. Khalidi 2010). I will not deal with this issue, but it does seem that the terminology of *interactive kinds* is more fitting for the phenomenon.

2.1 Real Kinds and Other Classifications

We may get a basic understanding of what categorisations in terms of *kinds* entail, by going back to how John Stuart Mill distinguishes between classifying in terms of Real Kinds and other types of classifications (cf. 1846, Book 1, Chapter VII, § 4). Let us compare two sets of statements:

A "... is an electron", "... is an acid", "... is a cat", "... is autistic"
B "... is white", "... is round", "... is obese"

The individuals categorised under the descriptions of group A has got *lots* of features in common with each other. This makes them useful for inductive inferences. They work in the way I described kinds earlier; as properties that cluster in the sense that some subset of properties are good indicators of the presence of the others. They do, potentially, give us a lot of information about the entities classified as such just by stating that an individual is a member of one particular such category.

When we compare this with the classifications in group B, we might say that there is some *kindhood entity* which the examples in group A represent, but which the ones in group B do not represent. These things have *one* feature in common—presented in their description—and from this we cannot infer many

useful things. We may say that there is *insufficient unity* to the entities grouped together based on this description, and the properties we may infer are in general quite trivial things, like the fact that round things tend to roll. Compare the statements about whiteness with statements regarding, for example, cats. If we know something is a cat, we will by this fact alone, know a whole range of things about the individual so classified— it's behaviour, what it likes to eat, how much it sleeps, how we should treat it, and so on.

2.2 A First Glance on Looping

As mentioned, an important part of the interactive kinds we find internally in the human species is that they are, to some degree, 'looping'. This notion entails that being classified as something will influence the people being classified, and that this, in turn, is reflected back on the classificatory process, perhaps to such a degree that the classifications themselves change. We may say that change in these kinds is the most central feature, but in special cases we may also have the *creation* and *destruction* of kinds.

Our example, autism, is a diagnosis which may be said to be constructed, in the sense that it was created in the 1940s as a category, and which has changed immensely since its introduction.[29] Regardless of this change, we hold it to be a real category, with a solid biological foundation. The category of autism

[29] It seems like too strong a statement to say that all changes in the autism diagnosis have come about because of looping. There are several factors at play in the short history of autism, but *part* of the reason why this category has changed so much is that there has been a feedback effect between the classification and the classified. This means that the classifications we use have changed as a response to how autistic people actually behave.

(as well as our understanding of autistic people) has been 'wandering' since the introduction of the diagnosis, and the definitions of the kind has changed enormously since its introduction. What counts as autism and what we expect from autistic people have changed, these changes are not only relevant for the classification in itself or for the diagnostic criteria, more importantly, we find quite substantial changes regarding the ways we talk about this group of people, how we address the people classified, and how the classified view themselves.

3. The Autism Diagnosis

To see how autism is suitable as an example when discussing looping, we need some information about the history of the autism diagnosis, from its dawn in the 1940s and onwards. The diagnosis has gone from being seen as a rare childhood condition to being considered a lifelong condition manifesting in many forms. The diagnosis is first described in 1943, then as Early Infantile Autism, by Leo Kanner (1943).[30] However, the condition was simultaneously described by Hans Asperger, whose work remained largely forgotten until re-introduced by Lorna Wing in the early 1980s (Asperger 1991; Wing 1981).[31]

By the time Hacking published 'Kinds of people: Moving targets', the understanding of autism had radically changed from those first descriptions. In 2007 the understanding of autism as a *spectrum* was commonplace, and we would be speaking of

[30] The use of the word *autism* is seen even earlier, but then primarily as a way of describing particular features of schizophrenia.

[31] My synopsis of the history of this diagnosis has been kept as short as possible. For a comprehensive, yet introductory, publication on the history of autism, see Silberman (2017). For a more specific introduction to the autism diagnosis, see Fletcher-Watson and Happé (2019).

conditions like *high-functioning autism* and *Asperger's syndrome* (hereafter AS). AS was introduced by (Wing 1981), but not seen as an official diagnosis until it was included in the DSM-IV.[32]

In general, autism has been understood as a *triad of impairments*. This is the understanding which was used in both the DSM-IV and the ICD-10, and it is the paradigm Hacking operates within. According to this paradigm, autism is seen as a combination of the following three traits:

1. Impairment in social interaction
2. Impairment in communication
3. Restricted interests and repetitive behaviour

Asperger's syndrome has been used to denote the people who have the symptoms of autism, but which lack problems acquiring language. This diagnosis is not present in the DSM-5. In addition, when the DSM-5 was released in 2013, the two first traits were combined into one, thus viewing social interaction and communication as more intertwined.

With these diagnostic criteria in place (the ones found in both the DSMIV and DSM-5), it is clear that autism is no longer a rare childhood condition, and with the inclusion of a larger autism spectrum the diversity among autistic people is substantial. With this in mind, we are ready to look at the phenomenon of *looping*.

[32] We can, of course, argue that the notion of autism as a spectrum was already introduced through Asperger's work, but since his work remained in obscurity, and was, for example, not available in an English translation until 1991, his syndrome was present in the sense that the knowledge *existed*, but it was overlooked.

4. Looping—A Framework

So far I have mentioned that there is a certain feedback between classification and the classified, between names and the people named. Hacking talks about how creating new names for things which did not earlier have a name is *in some sense* creating new things. However, we are not simply speaking of a relationship between the names and the things which are thus named in this case. There is a whole framework at play here, where several factors are contributing to what we shall call looping. All five elements, and probably more as well, are relevant parts of the interaction which gives rise to looping.

1. classification, or the naming of the phenomenon. In our case, we speak of autism, high-functioning autism, and AS, and we speak of *the autistic person*. This is the *kind of person* that is a moving target.

2. The people, in our case the autistic people themselves. In Hacking's example specifically the high-functioning autistic person.

3. Institutions, that is, clinics, meetings, and conferences. 'Institutions' is basically referring to *places where the diagnosis is discussed*, which also includes places where one can 'meet' the diagnosis. For some diagnoses, certain television programmes could be relevant,[33] and for autism, the Internet could be suggested as an additional institution where the diagnosis is met.

[33] Hacking suggests this to be the case for Multiple Personality Disorder (now known as Dissociative identity disorder), which he also discusses at length.

4. The knowledge, or the *presumptions that are taught* and refined, within the institutions, particularly what Hacking refers to as 'the basic facts'. He operates with two kinds of knowledge, fading into one another; the expert and the popular knowledge—the latter shared by a significant part of the interested population (Hacking 2007, 297).

5. the experts, or the producers of knowledge, refers to universities, hospitals and so on. The production of knowledge happens through several *engines of discovery*.[34] These engines are the means through which we gain new knowledge about a particular condition. For example, we may attempt to increase our knowledge about autism through figuring out things like the prevalence of autism, whether there is a genetic cause underlying most cases, whether it is better interpreted as a disability than a disorder, and so on, and through these engines our knowledge increases.

The idea is that as we get to know more about these properties which cluster to make a kind, we will be able to control, help, and also change these people. But the story turns out to be more complicated than that. Hacking argues that the people we are trying to know more about are *moving targets* because our investigations interact with the targets themselves, and change them. And because the people are changed, they are not quite the same kind of people as before; the target has moved. This

[34] I do not explicitly discuss these engines of discovery, but see Hacking (2006) and Hacking (2007) where this concept is introduced and explained in relation to the autism example.

is the looping effect. In addition, and importantly for our example, sometimes our sciences *create* kinds of people that *in a certain sense* did not exist before. Hacking calls this *making up people* (Hacking 2007, 293). I shall return to the idea of making up people in section 5.

Hacking's most used example, that of multiple personality disorder fits the framework very well, whereas the autism example is more complicated because of several factors, such as there being a pretty high chance that there is a strong biological component present in what causes autism.[35] However, while we are relatively certain there is a robust biological foundation here, the kind ('autistic people') has been wandering a great deal since it was first described in the forties. Being autistic now does not entail the same things as being autistic then. Another factor that complicates the picture when it comes to autism is that the condition *may be* inaccessible for the individual themselves. In some cases the autistic person will not have the knowledge that they are autistic, and in that case it is seen as an *inaccessible kind*.[36] I will briefly get back to the implications of autism as an inaccessible kind in section 4.2.

4.1 The Example of the High-Functioning Autist

Hacking considers the high-functioning autist (hereafter HFA) as an example described as follows: The HFA is, in Hackings terminology, someone who grew from an autistic child to an adult with full or almost full possession of language and 'some residual eccentricities of an autistic sort' (2007, 302).

[35] I will not go into the example of the looping effects regarding multiple personality here, but see Hacking (1995) and (1998) for more details.

[36] This is mainly the way Hacking uses the example in (1995).

First of all, we need to understand that it is not unproblematic to be talking about the high-functioning autist as a kind of person, because there are no clear limits to the categories of high- and low-functioning autism. This terminology is not particularly good, because there is little genuine information to be gained from it. Thus, we should ask *why* Hacking wants to use exactly the HFA as an example. Why is this specification important for him? That is, what does he think is entailed in the notion of the HFA as a kind of person which is no longer present if we speak of all autistic people instead? There must be something here, otherwise the specification is meaningless.

My interpretation of Hacking is that what he is looking for is *someone who are clearly aware that they are autistic, and can have reflections around this*. That is, people who are aware of (or potentially aware of) what this diagnosis means for themselves, what it means for the ones around them, and so on. Following this, I shall assume that he is using the HFA as an example in Hacking (2007) because he thinks it is important for looping to occur that the individual knows how they are being classified, and he is assuming that for this to be the case one needs to be, so-called, high-functioning.

Hacking suggests that the following scenario is a possible explanation for the idea of the HFA as a kind of person to arise: First these individuals had to be diagnosed with autism (of the first available Kanner description), and then they would have to 'miraculously recover' or 'grow out of it'. This entails things like acquiring social skills, understanding what people are thinking and feeling, and be able to live 'unproblematically with the obsessive need for literalness' (Hacking 2007, 303). Once these 'recovered' adults existed, other adults *not* diagnosed in childhood could be seen as having similar difficulties

'even if their childhood was not as bad' (ibid., 304).[37] I shall not discuss the historical accuracy of this story, but I argue that with some modifications this interpretation does capture at least some of the mechanics behind the expansion of the autism diagnosis. The modification which is needed has to do with the way Hacking is phrasing this narrative in terms of *recovery*. He is interpreting the first cases of high-functioning autism as something which came about because some autistic people went through a recovery process. However, it makes little sense to speak of *recovery* in the context of autism, because this implies that there is something to recover *from*.

However, even though Hacking's wording is inappropriate, what he seems to mean by a recovery is relevant for his storyline of the identification of the HFA. This means that we need to be slightly charitable in our reading of Hacking in order to get a grasp of what he is actually talking about. I contend that what is missing from Hacking's understanding of autism, which is also crucial to his idea of a recovery, is the fact that autistic people *develop*. Let us briefly clarify this. When describing change in an autistic person, for example the gaining of a skill that has been previously missing, this will often be presented as the person 'getting better' or recovering from autism. This completely fails to take into account the fact that autistic people can very well learn things and develop, and that this is not in conflict with being *just as autistic* as they were before gaining the skill in question. What might arguably be normal development of an autistic person is in Hacking's framework interpreted as *recovery*, and as something which is remarkable.

[37] A more correct way of phrasing this would perhaps be to say that their childhood was not *seen as being* as bad.

The description of the rise of the HFA as a kind of person is clearly Hacking's *interpretation* of what happened, when hitherto undiagnosed people were seen to exhibit the same traits as adults diagnosed with autism in childhood—and the subsequent expansion of diagnostic criteria. However, the core of the description seems correct. It is reasonable that one would be able to recognise autistic traits in other people based upon how autistic adults actually behaved, and following this, that more people would be seen to share the diagnosis. Thus, these other adults could recognise themselves as sharing the same condition, or clinicians could diagnose them as having the same condition. When someone is diagnosed with autism in adulthood we need to ask how we missed their autistic traits earlier. This means that we need to figure out which traits *these individuals* exhibited in childhood which would have made us able to diagnose this sooner; we look for holes in our classificatory system.

These mechanisms led to a growing number of adults identified as HFA, and after a while some of them would be classified as having AS. I am aware that Hacking wants to leave AS out of the debate, but given that the distinctions between the diagnostic criteria for Asperger's and autism are minuscule to such a degree that the former diagnosis is now rendered obsolete, I contend that this cannot be done. I do not know which assumptions is behind Hacking's exclusion of AS, but it may well be based on the common misunderstanding that the difference between AS and autism more generally is greater than it really is. In this case we should take into account the fact that Asperger's *is* autism, and that the only diagnostically relevant difference is that Asperger's excludes people with delays regarding the acquiring of language.[38]

[38] For more information about the distinction, or the lack thereof, between people considered to be HFA and people with AS, see Verhoeff

4.2 Looping in the Case of Autism

The main interesting moment regarding looping in the case of autism is when we start expanding who the autistic person can be, when we see that some people who were clearly autistic were capable of self-reflection, capable of acquiring language, and were very often not a child. This is something which is different from the early definition of Kanner's autism. The inclusion of high-functioning autism and AS into the idea of what autism is, alongside the idea of it as a spectrum marks very important changes. The inclusion of the so-called high-functioning people naturally lead to there being more people considered people *of that kind*, hence we get larger diversity in the kind, and hence also greater possibility of the kind changing as a result. These changes are *crucial* for the way we look at autism today.

Because of these changes, and the inclusion of more people and more possible clusters of properties, there is new knowledge to be had about the kind. As a result, the experts are forced to rethink their classifications, and so this new knowledge becomes part of *what is to be known* about members of the kind. *This* knowledge may again change the members of the kind, *or* lead to new people being included in or excluded from the kind. Both of these mechanisms lead to differences in the extension of the kind, and so the kind may change again. In relation to this, Hacking states that

> (t)o create new ways of classifying people is also to change how we can think of ourselves, to change our sense of self-worth, even how we remember our past. This in turn generates a looping effect, because people of the kind behave differently and so are different. (Hacking 1995, 369)

(2013). A discussion of the inclusion and exclusion of AS as a separate diagnosis is found in Howlin (2003).

In the case of autism, we have seen that this kind of self-reflection may be unavailable, because the individuals themselves are, in some cases, not aware that they are being classified as autistic. This is in line with how he presents autism in his (Hacking 1995), because here it is seen largely as an *inaccessible* kind. That is, the assumption was that these people—the autistic children—would have no knowledge of how they were being classified, and because of this it is much harder to see how looping will occur, as the self-reflection upon being classified as a particular kind of person is removed.[39] I do however argue that in this case we can maintain that there might be a reaction to being classified which is *independent* of the individuals knowledge of the classification.

The reaction will with great probability be *stronger* if there is knowledge about how one is classified, but we do not have reason to assume that looping does not occur in those cases where the classified individual is unaware of being so classified. Thus, we must be able to explain how looping can happen in the autism case, *without* the individual's knowledge of the categorisation being an essential part. There are, as far as I can tell, two options for how an autistic individual, who is known to be autistic, fails to know this themselves.

1. Either, the caregivers or the people who know are choosing not to tell the person, or they are postponing it, or

2. The autistic person is in such a state that they are not able to know that they are autistic—or, at least, they cannot know at present.

[39] Hacking speculates that looping may happen to some degree anyway, but even so he proceeds to consider solely the HFA (in 2007).

This excludes people who *are* autistic but where the condition is not known to *anyone*, such as very late diagnosed adults. The reason for this exclusion will be obvious. Such people may be treated as being *different* in many ways, but they will not be treated in the way specific to autism and autistic people. Not until they eventually are diagnosed.

The point I want to make is that the route through self-perception need not be the whole story. There will also be cases where the people themselves are not consciously employing a certain way of being—what we may call the autistic way of being—but where the people *around them* do. That is, the institutions and experts may think in terms of a certain way of being, and in the case of autism, so may parents or other relevant people, such as friends, teachers, and so on. What *passes for knowledge* of an autistic way of being in a sufficiently large portion of the population cause us to treat people belonging to a particular kind in a certain way, and through this the people may change irrespective of whether or not they themselves know the reason why they are treated in this particular way. If you are treated as being a certain kind of person, there are great chances you end up conforming to it, or at least that you react to it. This is of particular interest for autism because there is an amazing range of information which passes for knowledge in this area, but which are rather misconceptions and stereotypes which are reproduced as knowledge in society.[40] It would be more surprising if this did *not* influence people.

[40] For some of these stereotypes the representation of autism we find in fiction should be seen as a relevant source, see e.g. Nordahl-Hansen et al. (2018) who claim that portrayals of autistic characters in film and TV align unrealistically well with DSM-5 diagnostic criteria.

Hence, I maintain that the people classified will *most commonly* be actively involved in the process, but that we should consider cases where they are not. Provided that the rest of the framework for looping is intact it is likely that looping will happen also in these cases. There are no good reasons to restrict our example to that of the HFA.

So far we have seen that knowing that a person is autistic will make us treat the person differently—we now treat them according to our knowledge of this group of people—and this may change the person *even though* they do not themselves know they are classified in this particular way. However, for autism, just as it is possible that self-reflection *need not be* a necessary part of the mechanics of looping, the people themselves can also be *very actively involved*, because looping is not only something which happens because *someone else* treats you in a particular way. Knowing that you belong to a particular category can change you because you are given tools to understand *yourself* differently. This active attitude towards one's own categorisation is related to what Hacking refers to as 'resistance by the known to the knowers' (Hacking 2007, 306), and this resistance has been, and is, important for the development of how we view autism and autistic people themselves.

> Kinds of people who are marginalised, normalised administered, increasingly try to take back control from the experts and the institutions, sometimes by creating new experts, new institutions. (Hacking 2007, 311)

For autism this is seen in the *neurodiversity* movement, which is described by Hacking as a kind of 'autism liberation front'. Hacking's description of the movement is, however, too shallow and clichéd to be enlightening. He limits the movement's

significance to the people who can be described as high-functioning, and his description of it is as follows: "Stop trying to make us like you. We do some things better than you, and you do some things better than us, so leave us be" (ibid., 303).

Thus, his version of the neurodiversity movement amounts to much the same as the caricature where autistic people say that their autism is 'only a difference, not a disability'. We find this particular formulation in Miriam Solomon's article about the disappearance of AS as well, when she describes a debate over 'whether what is being diagnosed is a disability or *merely a difference* (conceptualised as "neurodiversity")' (2017, 11; my emphasis). But there is nothing in the idea of autism as a naturally occurring difference in neurotype which prevents it from also being a disability. Taking back the control over the knowledge of one's own group is not something which entails that the disability somehow magically disappears, and tying this control of knowledge solely to those individuals who may be seen as high-functioning is far too narrow an understanding. This movement needs a better philosophical understanding than the depiction provided by both Hacking and Solomon, because it is clear that autistic voices themselves have played an important part in reshaping of 'the basic facts' about autism.

5. End Remarks, and the Need for Further Research

So far, we have seen how looping work upon our classifications, and how kinds are being modified, how revised classifications are formed, and how 'the classified change again, loop upon loop' (Hacking 1995, 370). I have also briefly mentioned how the framework of looping can be utilised to speak of the *creation* of kinds. In a certain interpretation of 'creation' our example also shows this, with the first proper description of autism stemming from the 1940s. The autistic way of being was

not a way to be a person before the diagnosis was first described, the high-functioning autist was not a way to be a person until even later, and AS was not an official diagnosis until its inclusion in the DSM-IV in 1994, and the AS way of being a person was short-lived indeed (cf. Verhoeff 2013, 453).

Arguably, each of these subgroups of autism brought with them certain ways of being, but autism seems to be distinct from several other diagnoses in the DSM and the ICD diagnostic manuals, because it is generally seen to have a strong biological component. Hence, it seems odd, or even blatantly wrong, to say that there were no autistic people before the diagnosis was first described. There undoubtedly were autistic people existing in the world before 1943, but the particular autistic way of being came along first *after* the diagnosis was available. We may say that through the act of finding out the facts about autistic people we are also *creating* kinds of people, through constructing new *ways of being* which were not available earlier.

We have seen how autism works as an example of both the looping effect happening over time, and also of the creation of a kind, understood as a certain *way of being*, but interestingly we also find an example of the *destruction* of a kind within the autism spectrum. The creation and the destruction of AS as a separate diagnosis happened within very few years, something which is quite remarkable for a psychiatric diagnosis.[41] Solomon (2017) does address some of the more philosophical considerations regarding the removal of this diagnosis, but this is

[41] If we take the DSM system into account, it was included in 1994 and subsequently removed by the time the DSM-V was released in 2013. It is still possible to get an Asperger's diagnosis, provided you live in a country still using the ICD-10. However, the ICD-11 was released in 2018, and the same change is seen also here.

an area where more research is needed. A particularly interesting question is whether the AS diagnosis ever really denoted a particular *kind* of person with a particular *way of being*, another is what happens to this eventual kind when the diagnosis is removed. Such questions are outside the scope of this article, but I do contend that the case of the destruction of kinds which we see in the removal of AS as a separate diagnosis is one where the philosophers of psychiatry could make a valuable contribution.

We have also seen how the creation of new kinds of people may be said to happen through several *engines of discovery* in the human sciences, and even though we consider these to be ways of 'finding out the facts' about people, they are at the same time contributing to creating kinds of people that were not there before (Hacking 2007, 305). I contend that this is pointing towards a second area of research where the philosopher could be a relevant voice: We need to ask what information about autism we get through these different routes, and we should question both how this knowledge is interpreted, and what it can tell us about what autism *is*. To give an example, think about the research question about the causes of autism. What if we assume there to be some kind of genetic anomaly which explains all or most cases of autism. If we locate this genetic anomaly, would that entail that we know what autism is? Consider the following quote from Hacking.

> There may be, in the genetic make-up of human beings, a rather rare set of genetic anomalies that is responsible for most cases of autism. If so, it is a fixed target at which we aim, although we don't know what it is. The anomalies (if there are such) cause a person to be autistic. They do not determine the ways of life for autistic people (...) Think about the ways that the disability we call autism has changed its contours and its

> lived experience during the past sixty years. That is the moving target. (Hacking, 2006)[42]

If we agree with the point Hacking is making here, this entails that the knowledge of the cause of autism will not give us full knowledge of what autism is. Even more important, knowledge of the cause will not even entail that autism is static. The autistic people are not a static kind even if the kind has a solid biological foundation. This is in line with what the development of the autism diagnosis through the past decades have shown. The answers which may be gained through uncovering the cause(s) of autism are as such limited, and this should be clearly stated. Unravelling a biological cause will not tell us how the autistic *way of being* will be understood in the future, for example.

6. Conclusion

Adopting a wide and somewhat vague definition of kinds lets us speak of categorisations internally in the human species. This article has focused on a particular phenomenon occurring among such kinds, namely that of looping. I have shown how the autism diagnosis can be said to exemplify this phenomenon, and I suggest that there is no reason to limit this example to the so-called high-functioning autistic people (as is done by Hacking). Looping may occur also in cases where the individuals themselves have no access to knowledge of their diagnosis. Thus, the self-reflection appealed to by Hacking seems to be something which can intensify looping but not something which is needed for looping to occur. I have also aimed to highlight certain areas where further *philosophical* research is needed. This particularly concerns the interpretation of what knowledge is gained through the different investigations done

[42] This article is a slightly shorter version of Hacking (2007), and the cited part is not included in the latter version.

in medicine and psychology, and it also concerns the more particular case of the destruction of kinds.

Bibliography

Asperger, H. "'Autistic Psychopathy' in Childhood". U. Frith (ed.): *Autism and Asperger Syndrome*. Cambridge: Cambridge University Press, 1991, 37-92.

Bird, A. Essences and Natural Kinds. In *Routledge Companion to Metaphysics*. London: Routledge, 2009, 497-506.

Bird, A. and Tobin, E. "Natural Kinds". E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford: Stanford University, spring, 2018.

Ellis, B. *The Philosophy of Nature. A Guide to the New Essentialism*. Montreal: McGill-Queen's University Press, 2002.

Fletcher-Watson, S. and Happé, F. *Autism: A New Introduction to Psychological Theory and Current Debate*. London: Routledge, 2019.

Hacking, I. "The Looping Effects of Human Kinds". D. Sperber, D. Premack and A. Premack (eds.): *Causal Cognition*. Oxford: Oxford University Press, 1995, 351–394.

Hacking, I. *Rewriting the Soul: Multiple Personality and the Sciences of Memory*. Princeton: Princeton University Press, 1998.

Hacking, I. *The Social Construction of What?* Harvard: Harvard University Press, 1999.

Hacking, I. "Making Up People". *London Review of Books*, 28(16), 2006.

Hacking, I. "Kinds of People: Moving Targets". In *Proceedings of the British Academy*, volume 151, 2007, 285–318.

Howlin, P. "Outcome in High-Functioning Adults with Autism With and Without Early Language Delays: Implications for the Differentiation Between Autism and Asperger Syndrome". *Journal of Autism and Developmental Disorders*, 33(1), 2003, 3–13.

Kanner, L. “Autistic Disturbances of Affective Contact”. *Nervous Child*, 2(3), 1943, 217–250.

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., and Pellicano, E. “Which Terms Should Be Used to Describe Autism? Perspectives from the UK Autism Community”. *Autism: The International Journal of Research and Practice*, 20, 2015.

Khalidi, M. A. “Interactive Kinds”. *The British Journal for the Philosophy of Science*, 61(2), 2010, 335–360.

Mill, J. S. (1846). *A System of Logic, Ratiocinative and Inductive*. New York: Harper & Brothers, 1846.

Nordahl-Hansen, A., Tøndevold, M., and Fletcher-Watson, S. “Mental Health on Screen: A DSM-5 Dissection of Portrayals of Autism Spectrum Disorders in Film and TV”. *Psychiatry Research*, 262, 2018, 351–353.

Silberman, S. *Neurotribes: The Legacy of Autism and How to Think Smarter About People Who Think Differently*. London: Atlantic Books, 2017.

Solomon, M. “On the Appearance and Disappearance of Asperger’s Syndrome”. *Philosophical Issues in Psychiatry IV: Psychiatric Nosology*, 2017, 176–186.

Verhoeff, B. “Autism in Flux: A History of the Concept from Leo Kanner to DSM-5”. *History of Psychiatry*, 24(4), 2013, 442–458.

Wing, L. “Asperger’s Syndrome: A Clinical Account”. *Psychological Medicine*, 11(1), 1981, 115–129.

Nadia Mehdi

How to Tell the Stories of Others

For the last few decades there has been a mounting critique of the ways in which dominantly situated artists depict marginalised cultural groups, both historically and in the modern era. Some have argued that J. R. R Tolkein's depiction of the orcs in *The Lord of the Rings* belies belief in white racial superiority (Yatt 2002), whilst others have argued that Shakespeare's portrayal of Shylock in *The Merchant of Venice* has "an inherent potential for harm…[as a] symbol of Jewish vindictiveness, malice and hatred" (B'nai B'rith quoted in Sebag-Montefiore 2017). More recently, John Boyne's novel *My Brother's Name Is Jessica*, has been reproached for transphobia, a lack of authenticity, and repeatedly misgendering it's protagonist (Yossman 2019); and Jeannie Cummins has come under fire for appropriating the border-crossing immigrant experience in her novel, *American Dirt*, with critics arguing she has produced a work of inaccurate trauma porn (Zaragoza 2020). In the world of young adult fiction books are routinely pulled prior to release due to problematic portrayals of marginalised identities (Vartan 2019).

The debate is often polarised with one side arguing that artists should not represent that which lies beyond their own personal experiences, and the other side aiming to protect the romantic notion of artists as free to create whatever they please. The criticisms articulated above tend to lie on the former side of this debate, and centre around the problem of misrepresentation and the harms that it can cause to marginalised groups. Elsewhere, I provide a thorough articulation of what misrepresentations encompass and the harms that they cause to oppressed groups (Mehdi n.d.). I argue that the misrepresentation of the experiences and subjectivities of the oppressed by dominantly situated artists exacerbates and perpetuates oppression at both the material and ideological level. This is due to the capacity for misrepresentations to corrupt the social imagination, the pool of interpretive resources that a given social body share regarding concepts or understandings of identity. Importantly, under oppressive conditions, this pool tends to privilege dominant conceptions over those of the oppressed. Oppressed groups are more often than not misrepresented in the social imagination in ways that mystify the oppressive nature of social relations and lay the ideological groundwork for the preservation of oppression, and the dominantly situated are more often than not reliant on the social imagination for their conceptions of social reality. This entails that dominantly situated artists tend to lack accurate understandings regarding the what-it-is-like-ness of living under oppressive conditions, often leading to the misrepresentation of these experiences in artworks.

What, then, is to be done? Should dominant artists simply write what they know if they stand at a greater risk of creating misrepresentations, and avoid any inclusion of otherness or difference in their works? This is an unpalatable solution, and not

simply because censorship is an undesirable road to tread. Whilst misrepresentations can exacerbate oppressive social relations, their undesirability does not necessitate that anyone should be forcibly prevented from creating them. Moreover, legislating against such a practice would be incredibly difficult to police and potentially involve a bizarre prescription of correct and incorrect ways to represent. Additionally, realist works devoid of diverse casts or oppressive social structures are as misrepresentative as those that include diverse casts but misrepresent them. The different worlds to which we belong are not bounded. There is significant social overlap between cultures and other social groups such that "affirming the existence of incommunicability among cultures…presupposes adherence to a racialist, apartheid-like set of beliefs, postulating as it does insurmountable discontinuity within the human species" (Todorov and Mack 1986, 175). Oppressive histories and social structures are shared, even if our vantage points are different, with some more prone to an objective kind of truth, and others a partial kind, misconstruing social relations as fundamentally fair. Fiction that does not attend to this would only be a feigned attempt at reality and demonstrate a wilful blindness to those different from oneself. What's more, fictions can also allow us to develop empathy for others by granting us insight into different worlds and literary landscapes which fails to represent our diverse reality are often harmful in themselves. George Gerbner and Larry Gross (1976) have gone as far as to argue that "representation in the fictional world signifies social existence; absence means symbolic annihilation" (182). As Janisse Browning has put it,

> It's a constant struggle to develop a positive sense of identity in a world where you - as a Black person, and particularly, as a Black female - are either absent from most cultural production or are misrepresented as a racial or gendered anomaly...[my] existence in this country's [Canada's] recorded social memory has been represented only in relation to Eurocentric cultural perspectives (Browning 1992, 32).

Zoe Cunliffe (2019) has recently argued that diverse narrative fictions can work to counter and mitigate epistemic injustices through their potential to influence the social imagination, and therefore our shared concepts and notions of identity, for the better. And there exist instances in which dominantly positioned artists have engaged in subject appropriation and not misrepresented the subjectivities and experiences of the oppressed; Andre Aciman has been praised for his depiction of a queer summer romance in *Call Me By Your Name,* Leo Tolstoy for his depiction of women in *Anna Karenina* and Carson McCuller's for her depictions across abilities and sexualities, but perhaps most famously race in *The Heart Is A Lonely Hunter.* As such, we can clearly see that whilst fictive misrepresentations may often stem from a subject's privileged social positionality, they will not do so necessarily.

Given, then, that fiction (almost) fundamentally relies on the representation of otherness, misrepresentations are not inevitable, and diverse representation is desirable, we need solutions beyond the prescription that dominantly situated writers must refrain from representing that which they have not personally known or experienced. Whilst there are a number of contingent

reasons that misrepresentations are likely, we are creatures capable of change, capable of confronting social scripts and schemas that undergird oppression, capable of both highlighting these and various modes of resistance in art and imagining new worlds devoid of such pain. In this vein, I will present here three methods writers may use in order to avoid wrongful and harmful misrepresentations that align with the epistemic virtues of *humility*, *curiosity/diligence*, and *open-mindedness* (as characterised by Medina (2013, 42)).[43][44]

Responsible Research

The epistemic virtue of curiosity/diligence[45] involves a motivation to know what one cannot claim to already know, to fill in cognitive gaps. Creators of narrative fiction, of course, will already practice this virtue to some extent, researching modes of dress, eating habits, geographical features and so on, as well as other cultures and social groups. Yet the high incidence of misrepresentations and an overreliance on stereotypes in the works of dominantly situated authors suggests that further work needs to be done to diligently research the lives, experiences and subjectivities of oppressed groups to which artists do not belong. This will not be an easy task. Medina notes that under conditions of oppression "social arrangements and circumstances get

[43] Of course, the publishing, television and film industries have a part to play in rectifying the long history of misrepresentations and lack of diversity in narrative fiction, but due to limitations of space I will focus here on the actions artists can take.

[44] Are artists obliged to do this?

[45] Kate Wojtkiewicz (n.d.) in relation to how artists might create better representations characterises this virtue as *due diligence*, arguing that its practice is something that creators owe to their audiences.

in the way of these subjects doing the requisite work to achieve the relevant knowledge" (2013, 43) but, at the same time, a proliferation of resources exists to aid artists in their attempts to undertake responsible research.

Authors can of course travel to other communities to gather first-hand or testimonial research, but it is important to avoid what Nora Berenstain (2016) has termed *epistemic exploitation*, a process occurring "when privileged persons compel marginalised persons to educate them about the nature of their oppression…marked by unrecognised, uncompensated, emotionally taxing, coerced epistemic labour" (570). Moreover, this immersion must be undertaken honestly unlike, for instance, John Howard Griffith's *Black Like Me*, written following the author's donning of blackface and travelling through the civil rights era American South in order to publish a nonfiction work about the experience of racism (indistinguishable from what African Americans had already been saying). An artist is not a journalist, attempting to uncover important yet hidden truths for the public greater good. Such modes of research are predicated on an inappropriate voyeurism and are undertaken dishonestly. It is important to acknowledge that oppressed peoples may be wary of privileged interlopers due to the ways in which similarly situated peoples have previously misrepresented their communities. Janisse Browning reminds us that most marginalised communities "hold some things secret, some things sacred, and are wary of sharing too much of our knowledge because of past betrayals" (Browning 1992, 33). Trespassing into their realms regardless is to exercise a colonial kind of power.

It is still important, however, that this research is rooted in accounts of the experience of subjugation and oppression stemming *from* marginalised communities. Fortunately, there are non-exploitative avenues for research such as articles and online courses on writing accurate representations of all manner of marginalised and minority subjectivities and experiences created by people with non-dominant identities. The website *Writing the Other* runs courses on building inclusive worlds, writing asexual, native American, deaf and blind and transgender characters.[46] The writer Alice Slater has run a workshop entitled, 'How to Write Fabulous Fat Fiction' that aims "to consider the language we use to describe fatness, when and why we might choose to write a plus size character and how we can approach fatphobia in our writing without falling into the trap of being fatphobic ourselves," (Slater 2019). Moreover, authors can seek out venues in which the marginalised express themselves in an unmediated manner, and venues in which they critique misrepresentations. In order to write convincing black women, Kayla Ancrum suggests that writers spend some time reading literature written by black women for black women: "learning the way black women have discourse *among each other* is the first step to understanding their perspective AND emulating their voice" (2013). These venues will often contain a debunking of prevalent stereotypes. For instance, the writer Arie Farnam warns of stereotypes of disability such as "gracious suicides: disabled characters who nobly commit suicide to escape their unbearable existence and reduce the burdens on abled characters" and "fakers and manipulators:…[where] non-disabled people fake a disability and some

[46] See https://writingtheother.com/.

disabled people exaggerate their difficulties to get attention, sympathy or advantages" (2018). There is no shortage of autobiographical accounts of life with a subjugated social identity, and a multitude of historical, sociological and political works exist detailing why certain groups are oppressed, and how this oppression is invisibilised. Yet, whilst the inclusion of responsibly and respectfully researched diverse characters is important, many marginalised writers stress that the focus should remain on *character*. As Chella Ramanan puts it with regard to people of colour, writers should

> begin and end with character. Black and brown people are as nuanced as white people…Watch out for trying to overcompensate for past tropes by turning your characters into one dimensional, good people, who are passive and liked by everyone… Start by focusing on their character, not their race by making them interesting, giving them motivations and secrets and then bringing in elements of their culture or background, as a flavour, not their be all and end all. (Ramanan 2016)

Authors must remain aware that their research will contain contradictions. There will be no overarching unity in the self-representations of oppressed peoples, and this should not be surprising. As I have mentioned, oppressed groups are not homogenous. Yet, we might expect that there exists some unity in the criticism of stereotypes and controlling images, and some overlap in the articulation of the experience of oppression for how else would an oppressed group come to possess a group unity *as* an oppressed group? Moreover, care will need to be taken to not presume that one member of a group speaks for the group as a whole. Emmalon Davis (2016) argues that to do so would

involve an inappropriate inflation of credibility and potentially involve the harms of typecasting, compulsory representation and epistemic exploitation.

Attitude and Imagination

The epistemic virtue of open-mindedness involves the acknowledgement, respect and (to some extent) the inhabitation of alternative perspectives. Maria Lugones' (1987) notion of *world-travelling* provides a methodology for the cultivation of this virtue. For Lugones, a 'world' refers to a particular existence, or experience, the physical (or imagined) space occupied by persons sitting at a particular juncture in our social structure. Some worlds are more tenuously constructed than others, due to their relationship with other worlds. For instance, "A traditional Hispano construction of Northern New Mexican life is a "world", …[yet] in the face of a racist, ethnocentrist, money-centered anglo construction of Northern New Mexican life [it] is highly unstable" (10) due to the power inherent in the dominant construction of Northern New Mexican life, which threatens to annihilate the traditional construction. Shifting between 'worlds' is what Lugones refers to as 'travel' and is necessary for those who are not dominantly positioned, given that they will need to understand the dominant rules of the social structure to survive in it. But Lugones argues that world-travelling is possible, and important, for the dominantly positioned. Travelling to someone's world allows us to understand "what it is to be them and what it is to be ourselves in their eyes. Only when we have travelled to each other's "worlds" are we fully subjects to each other" (17). The knowing of others can occur in greater or lesser depth, but crucially, "without knowing the

other's "world", one does not know the other" (18). It is important that this understanding of 'what it is to be them' does not occur through the prism of one's already held beliefs, but rather that one fully and open-mindedly hears that Other in their own words. One border-crosses not in order to create one's accurate artistic representation, or in a spirit of paternalism, but in the name of regarding the Other as an equal, as devoid as possible of arrogant perception.

Over time, one can attempt to restructure one's mind through the conscious alteration of one's imaginative processes. Medina argues that the preventative work needed to minimise the mistreatment of marginalised and minority peoples exists not only at the level of practices, but also of the imagination, specifically "the circulation of ways of imagining collective subjectivities…that demean them and prevent their inclusion in the community or their equal standing within it" (2013, 250). He proposes that we attempt to radically alter our processes of dramatic imagination, examining the ease with which we imagine certain things – those things that align easily with the dominant techne's organisation of the world -, and the difficulty with which we imagine other things:

> our habitual and automatic ways of imagining need to be disrupted or transformed in order to repair or prevent unfair treatments – for example, if it has become habitual in a culture to represent women as weaker than men. The point here is to exert imaginative resistance where there is none (Medina 2013, 256).

The deliberate process of imaginative resistance exists at two levels, argues Medina; the object level, and the meta-level. Resistance at the object level is resistance exercised towards the imagining of certain things, for instance, that a gay man or a Latinx woman would behave in certain ways, or even that a white man acting in a certain way constitutes a certain kind of action. Kathryn Stockett's *The Help* has been widely criticised for its representation of the civil rights era Deep South. If Stockett were to have practiced the virtues of curiosity/diligence and open-mindedness, then eventually she may have been able to exert imaginative resistance that prevented her from inventing a stilted, child-like 'black' vernacular for her maids to talk in, or adhering so strongly to the mammie stereotype. Similarly, Stockett might have included a black man character that was not a wife-beater or an adulterer. Resistance at the meta-level, according to Medina, involves resistance to a whole perspective, usually ideological, such as white supremacy, ableism or the patriarchy. The film *Me Before You* has been accused of peddling an age-old 'better-dead-that-disabled' stereotype due to its disabled co-protagonist ending his life rather than living with a disability. If the films writers had exerted imaginative resistance at the meta-level then perhaps the film could have had a different ending, one that celebrates disability as merely a different form of embodiment, or reveals to audiences that disabled people are disabled by societies inadequate resources rather than their own bodies. Exerting imaginative resistance may not come easily at first, but becoming suspicious of the routes your imagination leads you down seems like a good place to begin when attempting to write better characters who are situated differently to yourself.

Ultimately, it seems that a proper cultivation of the virtue of open-mindedness will lead one "to understand that there are very powerful reasons for people to dispute your right to tell a story – reasons that stem from historical, political or social imbalances, you've already failed to understand the place and people who you purport to want to write about" (Kunzuru et. al. 2016). Lenore Keeshig-Tobias (2017) similarly argues that in order to tell the stories of (oppressed) Others, artists need to advance from a place of solidarity. The artist's interest in the Other should not begin and end with their story. A research process that is responsibly undertaken, alongside open-minded border-crossing will facilitate an understanding of the *reason* representations are so important, since they will facilitate a better understanding of alternate interpretive resources to those in the social imagination.

Changing the Status Quo

The epistemic virtue of humility involves the possession of a humble and self-questioning attitude to one's cognitive repertoire. Medina explicitly states that the development of this virtue can lead to many epistemic achievements including "identifying one's cognitive gaps and what it would take to fill them" (Medina 2013, 43). As argued above, one can of course come to fill one's cognitive gaps through responsible research into the experiences and subjectivities of the oppressed, but in some instances this will not be sufficient. As such, the artist fully devoted to developing her humility may seek to reject the paradigm of Romantic individualism that currently constitutes archetypal fiction creation and seek to involve various modes of (appropriately remunerated) collaboration in her artistic prac-

tice. This is especially important given that even the most committed anti-oppression advocates will find it hard to truly rid themselves once and for all of the oppressive schemas that structure the social imagination and therefore our shared concepts and understandings of identity. George Yancy (2008), for instance, writes that whiteness has an ambushing quality, since the subject position of white people is productive of a form of ignorance (what Charles Mills has called *white ignorance*). Racism is embedded within our embodied and habitual engagement with the social world, "weaved within the unconscious, impacting everyday mundane transactions" (Yancy 2008, 230). Even the most anti-racist campaigner may find themselves ambushed by whiteness and the baggage that comes with it. For instance, Yancy recalls a story from Tim Wise who he describes as "trapped….within the vortex of white power…whiteness waylays the white self even as one fights against racism with good intentions" (ibid.). Despite Wise's anti-racist work, upon boarding a plane with two black pilots he thought "Oh my God, can these guys fly the plane?" (ibid.). This kind of ambush occurs in fiction as well. The novelist Sarah Schulman writes that after meticulously researching the black character in her novel, she was informed by Jacqueline Woodson that the black co-protagonists 'discovery' that her grandfather had been married to a white woman ultimately put "white consciousness into the mind and mouth of a black character…[since] concern about hidden racial mixing was a white anxiety…[and] black people know the history of slavery and rape, and don't carry the same concepts of racial purity as white people" (Schulman 2016). Our experiences are shaped and constrained by the specificities of embodiment and the social space that we can inhabit

and imaginatively adopt. As such, the dominantly situated author must fully recognise the weight of the baggage that comes with their subject position. They must fully acknowledge that even when they think that they have done a good job in rendering a marginalised identity, there still may be work to do.

Collaboration can take any number of forms. The white male creators of *Tangerine,* a film whose protagonists are transgender sex workers of colour, began researching by spending months with real life transgender sex workers in Los Angeles to "get the vernacular right… because it's not respectful or responsible to do it any other way" (Sean Baker quoted in Buckmaster 2015). Yet Sean Baker, the projects director, also stated that "after a while you become friends. You become this fellow artist working on the same project with the same goal in mind. The barriers fall away" (ibid.), and one of the protagonists, Kitiana Kiki Rodriguez was closely involved in post-production, providing extensive notes and feedback.[47] Collaboration is not rare in film and television, and there is no reason that other narrative works, such as novels, need to be any different. [48] Additionally, writers can engage sensitivity readers: "part fact-checkers, part cultural ambassadors" (Waldman 2017) sensitivity readers employed to highlight misrepresentations in art works, primarily novels. They do not force writers into a pre-set arena of permissible representations but rather aid them

[47] It should be noted that Rodriguez did not receive writing, production or editing credits calling Baker's comments about the value *he* places onto collaboration into question.

[48] For instance see "Can You Write A Novel As A Group" (Dovey 2019), a *New Yorker* article about a number of groups which have written collaborative novels.

in representing otherness, either at the beginning of a writing process or as the project advances. Importantly, sensitivity readers are not a guarantee against making a mistake, nor a replacement for an artist conducting thorough and respectful research. They cannot comprehensively summarise marginalised group identity, or be used as a final authoritative representative voice. Representations that irk one reader may not appear irksome to another. It is especially important that even though feedback has been sought in order to practice humility, that feedback is also accepted with humility. As the novelist Naomi Alderman writes, "people criticising your writing is not a violation of your freedom of speech, it is a gift freely given and should be accepted with gratitude" (Kunzuru et. al. 2016).

Both these modes of collaboration capture the spirit of Jose Medina's (Medina 2013) polyphonic contextualism which contends that communicative contexts are always pluralistic. Whilst the social imagination may uphold various interpretations, concepts, schemas and understandings of identity, these are in no way exhaustive and oppressive folks will often possess alternative, and oppositional, interpretive resources with regards to social reality. As mentioned above, there are good reasons for oppressed groups not to part with these. But those that are willing to can aid the artist practicing humility through providing an interpretive counterpoint with regards to the experiences and subjectivities of the oppressed.

Conclusion

It is likely that no one of these proposals will be enough to free a writer from their standpoint, nor prevent problematic imaginings. Nor is it certain that these gestural solutions are enough.

However, they are a start. Misrepresentations are not harmless, even when fiction. If we are going to demand fair and 'truthful' representations of the oppressed, as appears to be increasingly the case, then social epistemology has some of the tools that artists will need in order to create these.

Bibliography

Ancrum, K. "How to Write Women of Colour and Men of Colour if you are White". *Media Diversified*, https://mediadiversified.org/2013/12/19/how-to-write-women-of-colour-and-men-of-colour-if-you-are-white/, 2013, accessed 10/04/2020.

Berenstain, N. "Epistemic Exploitation". *Ergo*, 3(22), 2016, 569–590.

Browning, J. "Self-determination and Cultural Appropriation". *Fuse Magazine*, 15(4), 1992, 31–35.

Buckmaster, L. "Tangerine director Sean Baker: isn't it time diversity was taken more seriously?" *The Guardian*, https://www.theguardian.com/film/2015/aug/11/tangerine-director-sean-baker-isnt-it-time-diversity-was-taken-more-seriously, 2015, accessed 20/04/2020.

Cunliffe, Z. "Narrative Fiction and Epistemic Injustice". *The Journal of Aesthetics and Art Criticism*, 77(2), 2019, 169–180.

Davis, E. "Typecasts, Tokens, and Spokespersons: A Case for Credibility Excess as Testimonial Injustice". *Hypatia*, 31(3), 2016, 485–501.

Dovey, C. "Can You Write A Novel As A Group?". *The New Yorker*, https://www.newyorker.com/books/page-turner/can-you-write-a-novel-as-a-group#:~:text=, 2019, accessed 05/06/2020.

Farnam, A. (2018). "Advanced Writing Tips: Problematic and irritating disability tropes to avoid in writing". *Arie Farnam By The Hearth*, http://www.ariefarnam.com/books/2018/2/16/problematic-and-irritating-disability-tropes-to-avoid-in-writing, 2018, accessed 10/04/2020.

Gerbner, G. and Gross, L. "Living with Television: The Violence Profile". *Journal of Communication*, 26(2), 1976, 172–199.

Keeshig-Tobias, L. “Stop stealing native stories”. *The Globe and Mail*, https://www.theglobeandmail.com/news/national/cultural-appropriation-stop-stealing-native-stories/article35066040/, 2017, 05/06/2020.

Kunzuru, H., Shamsie, K., and Alderman, N. “Whose life is it anyway? Novelists have their say on cultural appropriation”. *The Guardian*, https://www.theguardian.com/books/2016/oct/01/novelists-cultural-appropriation-literature-lionel-shriver, 2016, accessed 17/04/2020.

Lugones, M. “Playfulness, ‘World’-Travelling, and Loving Perception”. *Hypatia*, 2(2), 1978, 3–19.

Medina, J. *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and the Social Imagination*. Oxford: Oxford University Press, 2013.

Mehdi, N. *Telling The Stories of Others*, n.d.

Ramanan, C. “How to write people of colour and ‘the other’”. *Writer’s Anon*, https://writersanontaunton.wordpress.com/2016/05/25/how-to-write-people-of-colour-and-the-other/, 2016, accessed 10/04/2020.

Schulman, S. “White Writer”. *The New Yorker*, https://www.newyorker.com/culture/cultural-comment/white-writer, 2016, accessed 27/05/2020.

Sebag-Montefiore, C. “If a Shakespeare play is racist or antisemitic, is it OK to change the ending?”. *The Guardian*, https://www.theguardian.com/culture/2017/nov/03/if-a-shakespeare-play-is-racist-or-antisemitic-is-it-ok-to-change-the-ending, 2017, accessed 17/04/2020.

Slater, A. “How to Write Fabulous Fat Fiction”. *Eventbrite*, https://www.eventbrite.co.uk/e/how-to-write-fabulous-fat-fiction-with-alice-slater-tickets-75599593419#, 2019, accessed 10/04/2020.

Todorov, T. and Mack, L. "'Race', Writing, and Culture". *Critical Inquiry*, 13(1), 1986, 181–191.

Vartan, K. "A timeline of 2019's 5 biggest YA controversies". *Entertainment Weekly*, https://ew.com/books/2019/11/20/a-timeline-of-2019s-5-biggest-ya-controversies/, 2019, accessed 21/05/2020.

Waldman, K. "Is My Novel Offensive?" *Slate*, https://slate.com/culture/2017/02/how-sensitivity-readers-from-minority-groups-are-changing-the-book-publishing-ecosystem.html, 2017, accessed 20/04/2020.

Wojtkiewicz, K. *What We Owe To Our Audience: Creators of Fiction and Their Hermeneutical Responsibility*, n.d.

Yancy, G. "Whiteness as Ambush and the Transformative Power of Vigilance". *Black Bodies, White Gazes: The Continuing Significance of Race*. https://doi.org/10.1353/mel.0.0061, 2008.

Yatt, J. "Wraiths and Races". *The Guardian*, https://www.theguardian.com/books/2002/dec/02/jrr-tolkien.lordoftherings, 2002, accessed 17/04/2020.

Yossman, K. "Writers Blocked". *The Spectator*, https://www.spectator.co.uk/article/writers-blocked, 2019, accessed 17/04/2020.

Zaragoza, A. (2020). "The 'American Dirt' Controversy Illustrates the Media's Thirst for Immigrant Trauma Porn". *Vice*, https://www.vice.com/en_us/article/wxenzm/the-american-dirt-controversy-illustrates-the-medias-thirst-for-immigrant-trauma-porn, 2020, accessed 17/04/2020.

Caroline Stampone

The Relevance of Arendt`s Accounts of Evil and Refugeehood Today [49]

Introduction

In this paper I discuss why Arendt's accounts of evil and refugeehood are still relevant to help us understand the problems concerning especially vulnerable migrants (EVM hereafter) in its contemporary forms. 'Problems' in the plural because the expression 'the problem of EVM' refers both to the problem that EVM have to deal with and to the problem that they pose to states to solve[50]. By EVM I mean genuine refugees, illegal migrants, asylum seekers, refugees and internally displaced persons in camps. As highlighted by Gillian Brock (2020), the majority of forcefully displaced persons does not arrive in wealthy states, instead 90% of them stay in the Global South. Although that is the case, we live in times where the discourse that European states are too generous towards EVM is once again gaining prominence. In my view, Arendt's accounts of refugeehood and evil are especially relevant to help us critically discuss the contemporary responses of states and individuals to the problem of EVM. That is the case because Arendt's theory

[49] This article is a partial version of my Master thesis, which was delivered in September 2019 and has not been published anywhere.

[50] Natasha Saunders make this same point about the expression 'the refugee problem'. See Saunders (2019).

demystifies the myth of generosity and exposes the fact that in our world the problem of EVM is approached as a matter of luck and not as a matter of rights. In other words, the refugee system that we have in place today is unable to protect the big numbers of refugees[51], instead, it protects just a small fraction of refugees. From that, it does not follow that what is happening in practice is legally or morally acceptable. In my view, Arendt's sharp discussion of the refugee problem as experienced in Europe 80 years ago can be read as a call for the urgent participation of individuals in our common political world.

Many scholars have been applying Arendt's account of refugeehood in order to try to understand contemporary versions of the refugee problem[52]. In the literature on contemporary refugees Arendt's oeuvre is often approached in order to illustrate that refugees today are still stateless and rightless. Megan Bradley argues that to transfer Arendt's characterization of the refugee to the contemporary refugee crisis can do a disservice to the refugees themselves (Bradley 2014, 102). She argues that Arendt's descriptive account of refugeehood is dated, that is, it

[51] I consider my point to be true whether we consider refugees in a narrow or in a broad sense. In this paper I am often using the term refugee and the expression EVM as synonyms. That is, I am embracing a very broad account of refugeehood. The expressions 'the problem of EVM' and 'the refugee problem' are also used as synonyms here, they refer to the political problem concerning the protection of refugees in a broad sense, that is, the protection of genuine refugees and other especially vulnerable migrants.

[52] When Arendt approaches the refugee problem, her focus is not on refugees as a problem that states need to remedy. Instead she focuses on the problem that refugees are forced to deal with. Her focus is on how such a problem is experienced by refugees themselves. Moreover, she explains how in her view such a problem is intrinsically related to the fact that we organize the world in states that have an almost absolute right to exclude.

is about refugees created in Europe migrating to Europe or to the USA in the first half of the last century. Today, contrarily to Arendt's times, most refugees are created in the Global South and migrate to neighboring countries of their original lost homes, which are also in the Global South. Moreover, Arendt's account of refugeehood is intrinsically related to her critique of human rights. Arendt claimed that the mass migration of refugees in Europe 80 years ago showed that our human rights practice is inefficient and insufficient. Contrarily to Arendt's times, today human rights are the main language and tools that refugees have in order to survive. Bradley appeals to the fact that one of the main contemporary solutions for the refugee problem is repatriation. By doing so, she builds up her argument against the relevance of an Arendtian characterization of refugees. Bradley argues that in order to have the right to be repatriated one needs to belong to a state[53]. Because of that, she insists that we must see refugees as citizens of their original states. She further highlights that it is fundamental to recognize the refugees' right to be agents in the process of negotiating their own return to their states of origin (Bradley 2014)[54].

[53] The focus of this paper is not to discuss whether or not repatriation is a viable solution for the refugee problem. For my purposes it is enough to highlight that repatriation is a complex "solution" that has its limitations. Repatriation is only a morally acceptable solution for the refugee problem when it is in fact voluntarily and when the refugees indeed return to a save place where they can rebuild their lives with decency.

[54] One of the points that I want to clarify in this paper is that from the fact that Arendt characterizes refugees as stateless and rightless it does not follow that she assumes that they are powerless. On the contrary, Arendt like Bradley argues in favour of refugees becoming agents in the process of negotiating the solutions for their problems. The radical difference between their approaches does not concern the fact of whether refugees are able to empower themselves, but rather

In my view, although Bradley is right to highlight that Arendt's account of refugeehood is dated, and as such it cannot simply be applied to contemporary versions of the refugee problem, it does not follow from that that Arendt's account of refugeehood is not relevant today[55]. The aim of this paper is to explore other aspects of the relevance of Arendt's account of refugeehood that go beyond the characterization of genuine refugees as stateless persons[56]. Moreover, I argue that it is important to consider side by side the relevance of Arendt's accounts of refugeehood and evil. According to her, the way we deal with refugees in our statistic politics is in itself an evil. In my view, Arendt's account of refugeehood can help us understand what it means to be forced to exist as a refugee; while her account of evil invites us to think about why refugeehood became a mass phenomenon and what we can and should do in order to remedy such a problem.

Here I will focus on three important reasons why Arendt's accounts of refugeehood and evil are still relevant today. (1) First,

---

what they consider the necessary conditions to enable political participation.

[55] Bradley is careful to limit her critique to the uncritical applicability by contemporary philosophers of Arendt's characterization of refugeehood. According to Bradley own words: "To be sure, important aspects of Arendt's arguments continue to resonate. However, in this article I will endeavour to bring into focus some of the important ways and circumstances in which Arendt's depiction of refugees as stateless, rightless, 'scum of the earth' has become anachronistic, such that it does not clearly reflect some of the most crucial challenges faced by many current refugees" (Bradley 2014, 102).

[56] Along this paper I will use the expression 'genuine refugees' to mean or what David Miller considers to be those who should be counted as refugees, or to mean those that today are recognized as and enjoy the rights of refugees. That is, I am using this expression to refer to the 'lucky' ones that are granted asylum or are resettled. By 'genuine refugees' I do not include refugees in camps.

I will argue that to some extent the same main features that were true for the new type of refugee whose life conditions Arendt described are still true for non-genuine refugees in our days[57]. In both scenarios we are dealing with a problem in large scale; moreover, in both situations the right to asylum as a substitute to national law is not available to most EVM. (2) Second, I will argue that Arendt's explanation of why it was possible that in Europe 80 years ago the refugee question became an evil invite us to rethink how we justify the importance of refugeehood. Arendt argues that the phenomenon of statelessness is a by-product of our statistic politics, that is, the existence of large numbers of refugees is not a consequence of the failure of some states. Instead, it is intrinsically related to the fact that we live in a world that is organized in states that have an almost absolute right to exclude. I will situate Arendt's justification of why we should care about refugeehood in contraposition to the two main philosophical justifications for our obligations towards refugees. (3) Third, I will explain why Arendt's account of evil is relevant to help us reflect about the fulfillment of our obligations as individuals towards refugees. If related to her remarks about the banality of evil and about our duty to avoid evil, her account of statelessness as an evil can function as an invitation for *individuals* to rethink their own roles in remedying problems like the refugee problem[58].

---

[57] By non-genuine refugees I mean all those that are not lucky enough to be recognized as genuine refugees. I am referring to non-genuine refugees in this paper as EVM. By EVM I mean those migrants whose most basic human rights are unprotected, that is, refugees in camps, internally displaced persons, asylum seekers, and persons fleeing to escape extreme poverty.

[58] Note that I am not arguing that Arendt does not consider the role of the state in remedying the refugee problem. In my view, her take on responsibility reflects well the complex task of finding a balance between practical and ethical concerns regarding the roles of individuals

1.

At the core of my argumentation in favor of the relevance of Arendt's accounts of evil and refugeehood today is the thesis defended by Andrew Shacknove (1985) and Arendt herself according to which one of the ironies of our times is that to be recognized as an 'authentic' refugee (i.e. to be a refugee in a juridical sense) has become a kind of a privilege. That is the case because this category of vulnerable persons has real possibilities of enjoying at least some rights. As Arendt highlights, the same is not true for a new type of refugee that was created by our statistic politics and appear in masses. In the first part of this paper, I will consider Arendt's account of refugeehood as a synonym of statelessness and rightlessness. First, I highlight that Arendt's approach to refugeehood is of not much use to understand the political situation of those who have obtained

---

and the roles of states in dealing with political problems. On one hand, she insists that only individuals as part of a group can act to change the world. On the other hand, Arendt's insistence about the fundamental importance of personhood and of institutions to in fact protect the rights, the decency and the lives of human beings, reminds us that individuals alone are not able to protect refugees. States and its institutions are the main actors responsible for protecting all human beings, including refugees. However, in order to have access to this protection, one needs to *in fact belong* to a state or at least to a political community. Thus, in my view, the main difference between Bradley's and Arendt's approaches to refugeehood is not whether refugees *should* be seen as citizens of a state. They both recognize the fundamental importance of refugees becoming or remaining citizens of a state as an essential part of remedying the refugee problem. Moreover, they both recognize that political agency demands citizenship or at least membership in a political community. The main difference in their characterizations of the refugee problem lies in their different understandings regarding how one should take under consideration the limitations of the real world while theorizing about the refugee problem.

the status and the rights of authentic refugees[59]. By that, I mean the few EVM lucky enough to be recognized as refugees, and to receive protection of a hosting state in the form of asylum or resettlement[60]. In my view, however, those refugees who are forgotten in refugee camps should not count as authentic refugees. Today, only a very small proportion of refugees receive the legal right of authentic refugees. Arendt's account of the "new type of refugee" is useful to approach the problem of EVM precisely because it deals with those who are rightless in practice. I will show that according to Arendt the plight of EVM consists in being made superfluous, that is, expelled from the common world.

For Arendt, 'the authentic refugee' is happier than 'the new type of refugee', because the former has rights to protect her, such as rights to guarantee that she will find a new home, that she can be assimilated somewhere, and that somehow, she shares the common world with others. That is not to say that

[59] Arendt's remarks are still relevant for considering the existential aspect of the situation of the refugee in a broad sense, including the so called 'authentic' or 'genuine' refugees. One relevant critique in this regard – inspired by Arendt's approach to the problem of EVM – concerns the fact that such problems are often dealt with from the perspective of states, and not from the perspective of the refugees. For more about it see Saunders (2017).

[60] Bernstein refers to Arendt as one of these lucky refugees. He uses the expression 'lucky' in a broader sense though. He is referring also to the 'luck' through her whole trajectory as a stateless person. She was lucky to survive, given that there were so many situations in which she could have been killed. She was arrested in Germany for doing illegal research related to the situations of the Jews. After escaping to France, she was put in an internment camp. And other than many, she was able to survive, as well as her mother and her husband. Then after 18 years as a stateless person, Arendt finally was granted the right to nationality again (see Bernstein 2018, 3-8).

Arendt is uncritical about the status of the authentic refugee. She is an incisive critic of assimilation and made important remarks about how existentially damaging it is to be forced to exist only as a refugee, without being able to participate in the public space as *who* one distinctively is. To me, however, she does not claim that authentic refugees are necessarily rightless. In this paper, I am not approaching her account of the refugee problem in order to discuss the plight of those few who are recognized as –and enjoy the rights of– genuine refugees. Instead, I am approaching her characterization of the refugee in a broad sense, which she calls 'a new type of refugee', and defines as stateless and rightless (see Arendt 1962, 278-9) to discuss the plight of other EVM.

Although many scholars have been using Arendt to discuss the refugee problem in different contexts, it is relevant to keep in mind that a big part of Arendt's reflection on the refugee problem is a description of the situation of non-authentic refugees in Europe in the first half of the 20th century. By non-authentic refugees she means forced migrants who did not have the rights that authentic refugees have, nor any other rights, but instead were dependent on charity to survive.

One could object that Arendt's remarks aren't very useful to think about the refugee problem today since her focus was on refugees created in Europe and migrating to other European countries or to North America, while today the big majority of refugees are created in the Global South and end up seeking refuge in the Global South[61]. Perhaps this critique is relevant to those who use Arendt's account of refugeehood in order to think about the rights of genuine refugees, but that is not what

[61] As mentioned in the introduction such objection was presented by Bradley (2014).

I want to do here. In my view, what Arendt had to say about refugees decades ago is still relevant when considering the situation of EVM today. More importantly, I think it is relevant to approach the plight of EVM through Arendt's theory because she discusses the phenomenon of statelessness "not as an aberrant or accidental phenomenon occurring despite the best efforts of states to prevent it, [but as] a normalized systemic condition produced by an international order predicated upon the sovereign power to exclude as the essence of statist politics" (Hayden 2010, 457). That is to say, she explains how the problem of EVM is essentially related to the way we organize our world.

Moreover, one could still object the use of Arendt's account of refugeehood to think about the problem of EVM today because of the fact that human rights as a practice has improved significantly in the last decades. It is true that since the time that Arendt wrote there have been "significant legal, political, and normative changes" in our human rights practice (Gündoğdu 2015, 5). Perhaps the most important was "the international institutionalization of human rights norms" (Gündoğdu 2015, 5-6). Contrarily to the times when Arendt wrote, today we do have some international guarantees of legal personhood, such as the two binding international covenants adopted by the UN in 1966 and in effect since 1976. Moreover, the status of human rights is also quite different today from what it was 80 years ago. Human rights are no longer a kind of stepchild that no political party takes seriously (see Arendt in Gündoğdu 2015, 7), instead, human rights are a very important part of the agenda of many NGOs and it is incorporated into the foreign policy of many states (see Gündoğdu 2015, 7).

Although human rights as a practice has improved in the last decades, we are still treating refugees and other vulnerable migrants in an unjust manner. Today, as highlighted by Gündoğdu, the formal enforcement of human rights continues to be weak in the international level (Gündoğdu 2015, 7). Moreover, despite the developments regarding human rights, the condition of EVM "remains in that murky domain between legality and illegality" (Benhabib 2004, 154). Currently in Europe, we often hear that European states are being too generous in their treatment of refugees. There is a tendency to treat the problem of the protection of EVM as a matter of charity, instead of a matter of rights. That is one of the reasons why Arendt's characterization of the refugee as stateless and rightless is still relevant today. However, we also need to be careful about how we use Arendt's notion of refugeehood in our efforts to understand a contemporary problem. I consider Gündoğdu's reinterpretation of the Arendtian concept of rightlessness very relevant to critically approach the problem of EVM today. For Gündoğdu we should apply it to contemporary problems "not as the absolute loss of rights but instead as a fundamental condition denoting the *precarious* legal, political, and human standing of [vulnerable] migrants" (Gündoğdu 2015, 93).

A New Type of Refugee vs. Authentic Refugees

The contrast that Arendt draws between a new type of refugee that appeared after the two world wars in the beginning of last century in Europe and 'its happier predecessor' illustrates well why the so called 'genuine' or 'authentic' refugee is the most privileged among the especially vulnerable persons. David Miller argues that we should count as 'genuine refugees' only those whose basic human rights are threatened and cannot be helped where they are (see Miller 2016, 77). For him, all other persons below the threshold of vulnerability should not be counted as refugees, as in theory they could have been helped where they

are. Shacknove disagrees and argues that from a philosophical perspective we should consider all especially vulnerable persons that depend on international protection to have their basic needs addressed as refugees (cf. 1985). Shacknove highlights that "ironically, for many persons on the brink of disaster, *refugee status is a privileged position*. In contrast to other destitute people, the refugee is eligible for many forms of international assistance, including material relief, asylum, and permanent resettlement" (Shacknove 1985, 276).

By contrasting the new type of refugee (who is rightless) to the authentic refugee (who has rights because she possesses personhood), Arendt explains two important features of the new refugee. First, contrarily to those that used to be called refugees, the new type of refugee is a part of a massive phenomenon and does not have access to the right of asylum. In Arendt's words:

> Civil wars which ushered in and spread over the twenty years of uneasy peace were not only bloodier and more cruel than all their predecessors; they were followed by *migrations of groups who, unlike their happier predecessors* in the religious wars, *were welcomed nowhere and could be assimilated nowhere*. Once they had left their homeland, they remained *homeless*, once they had left their state they became *stateless*, once they had been deprived of their human rights they were *rightless, the scum of the earth*. (Arendt 1948, 341, emphasis added)

Arendt compares the old type of refugee to the new one claiming that the former was 'happier' because although she was also in a vulnerable situation consequent of the loss of her home and of her political community, she had a realistic chance of finding a new home and a new state, and that was the case because the right to asylum was still available to her. In another passage, Arendt mentions that the old type of refugee was "driven to

seek refuge because of some *act committed* or some *political opinion held*" (Arendt 1996, 110, emphasis added). For Arendt, the possibility of finding a new home and a new governments' protection, as well as the possibility of acting and participating in politics in the first place, are consequences of the fact that the old type of refugee had personhood. That is, she was not only a human being and nothing else, she had "the *artificial* mask provided by the law" (Arendt in Gündoğdu 2015, 92). For Arendt it is the possession and enjoyment of such artificial mask that gives someone the right to be treated as an equal in the eyes of the law; and enables someone to participate in politics and as such act and have opinions. Given that, for Arendt, the possession of personhood is essential for a human being to have a chance to flourish.

The new refugee is destitute of this artificial mask provided by the law, and because of this, even before being made a refugee, she was fundamentally an *innocent* victim, unable to act or to have an opinion. Arendt emphasizes that innocence, as a synonym of a "complete lack of responsibility" was the mark of the "rightless [of the new type of refugee] as it was the seal of their loss of political status" (Arendt 1962, 296). The authentic refugee, to which she refers also as the 'political refugee' was responsible for her own acts, in the sense, that at least she had the right to act and thus the opportunity to be responsible. The same was not true for the new refugee, which "committed no acts and (…) never dreamt of having any radical opinion" (Arendt 1996, 110).

Arendt also distinguishes the new type of refugee from the political refugee still present in Europe 80 years ago. Once more, her first move is to highlight that contrary to the new type of refugee, political refugees are "of necessity few in numbers" and then acknowledges that those few authentic refugees "still

enjoy the right of asylum in many countries, and this right acts, in an informal way, as a genuine substitute for national law" (Arendt 1962, 295). She insists that the fate of these few political refugees have nothing to do with human rights being efficient. They weren't deprived of legality because human rights weren't able to protect them, but because some states still recognized their right to asylum, not as a matter of human rights, but as a matter of old custom.

In my opinion, Arendt's account of refugeehood, that is, her characterization of this new type of refugee, touches the core of the problem of EVM today. Once more, the refugee question is a matter of millions of persons, and once more, their protection is not approached as a matter of rights, but instead as a matter of luck. By that, I mean that some refugees will be recognized as genuine refugees and they will be treated accordingly, that is, they will enjoy their right to asylum or to resettlement. But the fact that they will be treated in the right way is not a consequence of us having a system in place that efficiently protect their rights. Instead, it is a matter of luck. Thus, they are still rightless, in the sense that most of them will depend on charity and will be subjected to arbitrary decisions in camps and detention centers. Next, I will reconstruct Arendt's critique of human rights, which is intrinsically related to her concept of rightlessness and her justification of why we should care about refugeehood.

2.

We just saw that one important reason why Arendt's account of refugeehood is still relevant today, is that, it is very pertinent to help us understand how unjustly we still treat most especially vulnerable migrants. That is the case, because the living condi-

tions of most EVM today are quite similar to the living conditions of the new type of refugee that Arendt was talking about 80 years ago.

Secondly, Arendt's account of refugeehood is also relevant because it invites us to question how we justify refugeehood. Why should we care about protecting refugees? Philosophically, we either answer this question by appealing to the justification of the state as a viable political unity that is able to offer security to its citizens; or by appealing to the moral argumentation according to which it is wrong to let human beings suffer when we are able to do something to remedy or avoid their suffer at little cost to ourselves.

The philosophical justifications that we commonly offer for caring about the refugee problem are related to the justification of our way of organizing the world in states; or to the simply moral consideration that the right thing to do when someone is in dire need is to help (if one has the capacity to do so). Different justifications of why refugeehood matters often entails in different accounts of who should be counted as a refugee. Next, I will briefly contrapose Andrew Shacknove's and David Miller's justifications and definitions of refugeehood.

Shacknove (1985) deals with his definition of refugeehood without worrying about any feasibility constraints. Instead, he focusses exclusively on the thought exercise that functions as our justification for protecting refugees in the first place. He asks why persecution was a sufficient cause to give someone the right to immigrate and concludes that the protection of refugees is justified by the fact that we value "the normal social bond" that we establish with our states. Persecution is just one of the numerous ways in which such a social bond can be severed (see Shacknove 1985, 277). By the 'normal social bond'

he means what ties citizens to a state or a civil society. Shacknove argues that "the primary purpose of civil society is to reduce each person's vulnerability to every other" (1985, 278). He takes for granted the widely accepted idea introduced in political philosophy by Hobbes that justifies the existence of a state as a way to offer some minimum security to human beings. In his famous *Leviathan*, Hobbes did the thought exercise of imagining how a society without a state would be. He arrived at the conclusion that in such a society, men would always live in fear of each other and only the strongest would survive[62]. Because of this generalized condition of insecurity, human beings would have accepted to open hand of some of their freedom in exchange for security, and that was how the state was created. That is to say, an essential part of the purpose of the state is to protect at least the most basic rights of its citizens. Or as Shacknove phrases it "to be minimally legitimate and tolerable, the commonwealth must reduce the citizen's vulnerability to others, all others" (1985, 279).

Thus, Shacknove's move in order to find a definition of refugeehood is first to ask: Why is a definition of refugeehood necessary in the first place? His answer is that it is of fundamental importance for human beings living together in organized societies that the social bond between themselves and their political

[62] Miller argues that although Hobbes might have been wrong about human nature, his point regarding us needing a political authority such as the state is still correct, and that is so because of a matter of trust. In Miller's words "Hobbes' real point is that in the climate of fear that would follow the breakdown of authority, the kinder, more trusting, side of human nature would be obliterated. And from what we know of human behaviour when people are caught up in civil war and other situations in which their very survival is at stake, he seems to have been right. We need political authority, then, because it gives us the security that allows us to trust other people." (Miller 200, 23).

unities will be protected. In cases where such social bonds are broken, it is of the interest of other states – and of the citizens of those states – that measures will be taken to protect those failed by their own states. Otherwise, the credibility of states as a useful political unity could be questioned. As highlighted by Shacknove the "absence of state protection" is at the same time what "constitutes the full and complete negation of society and the basis of refugeehood" (1985, 277).

Miller also acknowledges that part of the philosophical justification of why refugeehood matters to us is related to the states' and citizens' interest in protecting the idea that states are a useful political unity for all of us. However, this is not the focus of his argumentation regarding why states have an obligation of justice to help refugees. He argues that such *remedial* obligation[63] exists simply because when there are persons in need of help (that otherwise would not survive or would be seriously harmed) those with the capacities to help, have an obligation to do so. But this duty is limited. Capable states are morally allowed to consider costs and are only required to contribute with their fair share for the solution of the problem of EVM. I am mentioning all of this here, because Miller, while thinking about a morally acceptable definition of refugeehood, is considering all of this as well. While figuring out a definition of refugeehood, Miller does not only ask 'why some persons should have the moral right to be protected through refuge?', but at the same time he asks, 'how many vulnerable outsiders capable states have an obligation to protect?'. And given that according to his contextual theory of justice states are only re-

[63] As mentioned in chapter 1, for Miller a remedial responsibility "is a responsibility that falls upon all states able to provide the necessary refuge" (2015, 4). Moreover, he argues that remedial obligations "are limited by considerations of cost" (2016, 92).

quired by justice to contribute with their 'fair share' for the protection of refugees, Miller also asks 'How many would they be willing to protect?'.

I believe that Miller's intention is to come up with a definition of refugeehood that would be able to do justice to 'genuine refugees' as well as to the states that have the obligation to protect them. That is, he aims at a definition that would take the perspective of refugees' seriously and still would be able to function as a useful criterion for states to select among immigrants without being overwhelmed by too many refugees. In my view, Miller's definition of refugeehood prioritizes the interests of the states and is therefore unable to serve as a morally acceptable tool for states to select among immigrants. It excludes many categories of EVM, such as persons forgotten in refugee camps, persons fleeing extreme poverty and internally displaced persons in camps. In my view, Shacknove's definition of refugeehood is the morally acceptable one, the one that can be useful for us to think about what can be done in order to avoid the evil of statelessness in this world. After all, in order to be able to face a political problem in the most efficient way we first need to understand what the problem is. In my view, Shacknove's definition of refugeehood helps us to do this.

Arendt is not interested in understanding philosophically how to justify or define refugeehood. Instead she engages in a descriptive account of refugeehood that appeals to the notion of personhood in opposition to the notion of what does it mean to be/became simply a naked human being and nothing else. She concludes that to find oneself in the situation where one is only a bare human being is dangerous. While reflecting about the situation of refugees in Europe 80 years ago, Arendt realizes that those that lost the protection of their states, those that lost

their civic rights, automatically also lost their supposedly sacred human rights.

Human Rights and the Condition of Rightlessness

According to Arendt, human rights – as they existed in the last century and to a large extent as they still exist today – are inefficient. By that, she means that in practice human rights in a legal or institutionalized sense fail. As we will see, in this descriptive part of her critique of human rights, she will attack the massive failure of human rights by pointing out the fact that in practice many people are rightless[64]. She claims that exactly when people needed human rights the most, such rights were absent[65]. That is what she will call the paradox of human rights[66]. In Arendt's words:

> If a human being loses his political status, he should, according to the implications of *the inborn and inalienable* rights of man, come under exactly the situation for which the declarations of such general rights provided. Actually, the opposite is the case. It seems that a man who is nothing but a man has lost the very

[64] It is important to have in mind, that in the descriptive part of her critique of human rights Arendt is not talking about human rights as moral rights, contrarily to Miller, who holds on to this meaning of human rights.

[65] Arendt is referring among other facts to the persecution of the Jews by the Nazis.

[66] Arendt does not believe that it is possible to solve such paradox once and for all. However, she believes that we can improve human rights considerably if we understand better what they are in practice and what they should be in order to address our modern needs. She will argue that in Modernity human rights are a necessity to protect human dignity. And she will suggest that we approach human rights as the right to have rights, which Gündoğdu reads as Arendt's efforts to rethink human rights "in terms of a right to citizenship and humanity" (Gündoğdu 2015, 5).

> qualities which make it possible for other people to treat him as a fellow-man. (Arendt 1962, 300)

According to Arendt, the paradox of human rights consists in the fact that although human rights exist in order to protect human beings as human beings (and not as citizens), in practice, human rights were unable to protect those who lost the protection of their states. In her view this shows that we are dealing with human rights in the wrong way. Arendt's project is not only to point out what we are doing wrong, but also to think about how human rights could be made efficient.

Since the 18th century[67] we have accepted that human rights are the rights that human beings have by virtue of being human. That is the assumption present in all relevant documents to the practice of human rights today, such as the "United Nations Universal Declaration of Human Rights 1948", the "International Covenant on Civil and Political Rights" and the "International Covenant on Economic, Social and Cultural Rights". All those documents assume that human rights are a priori, in the sense that they are self-evident and independent of citizenship or any other type of membership in a political community. They also assume that human rights are essential, because they are grounded on human nature[68]. Arendt rejects both assumptions, because she refuses the idea that human rights are grounded in *the metaphysical idea* that we are entitled to rights simply because we are humans.

[67] What we today call human rights were first formulated as "The Rights of Man" and in that sense were already present in the French and American Declarations.

[68] Other essentialists accounts of human rights hold that they are grounded on morality.

According to Arendt, historical facts had shown that what we have been calling 'human rights' are instead civic rights. By that she means that a person only has a right while she has the protection of some kind of organized community[69]. To have a right, in this sense, means not only to possess a right in theory, but also to be able to enjoy such a right. According to Arendt, the appearance of the phenomena of mass migration of refugees made clear that a stateless person was also a rightless person. That is, it made visible that human rights understood as universal and independent of states, did not really exist. What existed were civic rights, that is, rights that one possessed and enjoyed by virtue of being a citizen. Arendt argues that the phenomenon of the mass migration of refugees was what made clear that human rights as approached 80 years ago were inefficient and insufficient.

For Arendt, the new refugee, namely, the refugee that was persecuted not because of something that she said or did, but instead, because she was part of a group (as in her case part of the group of Jews), had no rights. Contrarily to the 'authentic political refugee', the new refugee has no one and no law to protect her. A rightless person is someone who has been deprived of a place in the world and of a recognizable identity. On Arendt's words:

> *The prolongation of their lives is due to charity and not to rights*, for *no law* exists which could force the nations to feed them; their freedom of movement, if they have it at all, gives them *no right to residence* which even the jailed criminal enjoys as a matter of course; and their *freedom of opinion* is a

[69] Notice that for Arendt that is not to say that we need states as we know it in order to enjoy human rights. She simply highlights that we need to belong to some kind of organized community. Moreover, Arendt also highlights that nation states creates refugees.

> fool's freedom, for nothing their think matters anyhow. (Arendt 1962, 296, emphasis added)

In the passage mentioned above, Arendt mentions rights such as the right to food, the right to freedom of movement, the right to residence and the right to freedom of opinion. All of those rights are considered human rights according to the UN Declaration of Human Rights, which was adopted in December 1948[70]. As I read Arendt, they *should* be human rights, but in practice they are not, because they depend on states in order to be enforced. It is worth clarifying that Arendt published the work where she first presents her critique of human rights, i.e. *The Origins of Totalitarianism* in 1951. So, she was aware of the existence of the UN Declaration. However, the appearance of such document – which by the way is a non-binding document – did not change Arendt's mind[71]. That is the case in part because of the language adopted in the Declaration. In its preamble, it is stated that "recognition of *the inherent dignity* and of the *equal and inalienable* rights of all members of the human family is the foundation of freedom, justice and peace in the world" (UN, "Universal Declaration of Human Rights"). Article 1 claims that "all human beings are born free and equal in dignity and rights (…)" (ibid.). For Arendt, it is not true that we have inalienable rights, nor is it true that such rights are the

[70] In the UN Declaration the right to food is included in article 25, which says that "Everyone has the right to a standard of living adequate to the health and wellbeing of himself and of his family, including food (…)"; the rights to freedom of movement and the right to residence are mentioned in article 13; the right to freedom of opinion is mentioned on article 19 (UN, *Universal Declaration of Human Rights*).

[71] Arendt wrote an article criticizing the UN Declaration 'lack of reality'. See Arendt (1949, 37).

foundation of freedom or justice or peace, nor that we are born equal.

For her, the rightless people created by the Nazi regime showed the world that there are no inalienable and universal rights, neither is it true that all human beings are born equal. Arendt insists that in practice rightless persons are not equals to citizens. Unlike citizens, a rightless person cannot count on the law or the state to protect her. Instead, she depends on charity to survive. It follows that for Arendt it is wrong to insist that a refugee without protection, or an illegal migrant, or an extreme poor person really possess human rights[72]. For Arendt, in order for someone to have human rights in the sense of "enjoying" them, it is not enough to have them written on a non-binding piece of paper. Those rights need to be enforceable.[73]

From that, it does not follow that we should abandon human rights and focus exclusively on civic rights. According to Arendt that would not be enough, because one can in theory have civic rights and still be rightless. To be clear, Arendt is talking about civic rights as a synonym of human rights as we have them, in their insufficient form. To illustrate her point, she

[72] The inclusion of extreme poor people as rightless in an Arendtian sense is not uncontroversial, because in theory they do have the protection of a state and the rights of citizenship. However, in practice they are very often unable to enjoy such rights. For an account of extreme poor persons as rightless see Hayden (2010).

[73] According to Beth Singer "to say that you have a right that is not operative in a community is to say that you believe that you and other *ought* to have this right. In a situation of rightlessness, it's not that we have rights, but they are unprotected; rather, it's that we *ought* to have them, since they are a condition for human action; they *ought* to be made operative through human determination and institutions" (in Parekh 2008, 140 f.)

compares the condition of the stateless with the conditions of a soldier that dies in war, and of a criminal that is in prison. She claims that the stateless is absolutely rightless, but the same neither holds for the soldier who is deprived of his right to life, nor for the criminal[74], who is deprived of his right to freedom. She claims that neither the soldier nor the criminal have lost their human rights. On the contrary, a stateless person, which in theory have the rights to freedom of movement, the right to opinion, and the right to equality, has in practice lost all these rights, because without the right to belong to a political community, none of those other rights have any meaning (Arendt 1962, 296). By that she means that there is something more fundamental than the rights of citizenship. In her view, it is the right to have rights, which is what makes possible and give real meaning to all the other rights. Arendt argues that human rights are important, and in a sense more fundamental than civic rights. We need human rights, and because of that we need to figure out a way of making them efficient.

Another important aspect of Arendt's critique of human rights consists in understanding the circumstances in which they appeared as inefficient and insufficient, that is, during the decline of the nation states in Europe. Next, I will reconstruct Arendt's remarks on the relation between human rights and the nation state[75].

---

[74] As pointed out by Gündoğdu (2015), Arendt ignores the fact that many criminals in jails around the world do not have real access to the law. One example of it are persons that after finishing their time in jail continue incarcerated in Brazilian jails.

[75] According to Bradley in order to engage critically with Arendt's ideas about refugeehood one must be aware that they "are grounded in her historical analysis of the nature of the European nation state and the rise of totalitarianism" (Bradley 2014, 101). Although I agree with Bradley, I also want to emphasize that Arendt does more than

Nation State, Human Rights and Rightlessness

For Arendt after the First World War, with the 'rebuilding' of Europe through the creation of imposed nation states "the very phrase '*human rights*' *became* to all concerned – victims, persecutors, and onlookers alike – the evidence of *hopeless idealism or* fumbling feeble-minded *hypocrisy*" (1962, 269; emphasis added). She explains that human rights *became* no more than hopeless idealism or hypocrisy because of the ascendency of nationalism. Arendt explains that the "very conditions for the rise of nation states" are "homogeneity of population and rootedness in the soil" (1948, 345). The imposition of nation states that followed the *Peace Treaties* were not able to fulfill such conditions. As a consequence of it, 30 per cent of Europe's population at that moment were made of recognized minorities, which as refugees, had the options of living under "the law of exception of the Minority Treaties" or "under conditions of absolute lawlessness". (Arendt 1962, 269).

She argues that in such scenario human rights became inefficient and were not more than an unrealistic idea. That was the case because human rights depend on political communities to enforce them. The problem was that the type of political unity of that moment, that is, the nation state, is the type of political organization that only takes care of those that are its citizens, leaving unprotected all those that are not part of the homogeneous population. Contrarily to the state, whose "supreme function" was to protect the human rights of "all inhabitants in its territory no matter what their nationality", the function of the nation state is "to distinguish between nationals and non-nationals and on this basis "grant full civil and political rights only

analyzing this specific historical circumstance. I consider her reflections about the differences between nation and state still quite relevant today.

to those who belonged to the national community by right of origin and fact of birth" (Arendt 1948, 296 in Hayden 2008, 251f.).

Arendt contraposes the state to the nation state and claims that the former is an instrument of the law committed to the protection of the rights of all human beings in its territory and compromised in considering all its inhabitants equals in front of the law. The same is not true for nation states, given that they are moved by this story about the great value of a nation, they are committed to the protection of only those that are full members of the nation. Moreover, in this type of political community only nationals are equal in front of the law. According to Arendt, human rights were idealized in a way that although they were supposed to protect the rights of human beings as human beings, they depend on states to enforce them. This tension at the core of human rights is a consequence of an approach to such rights that at the same time tries to reconcile the sovereignty of the individual with the sovereignty of the nation[76]. For Arendt, the phenomenon of mass statelessness reveals this tension, that is, "the destructive contradiction between universal human rights and the sovereign power of the modern state" (Hayden 2008, 249).

Such contradiction is destructive because it is responsible for the creation of the evil statelessness. That is, it is responsible for making human beings as human beings superfluous. That is the case because exclusion is an intrinsic part of how the European modern nation state functions. Stateless persons or refugees are simply the byproduct of the common practice of ex-

[76] This tension is also present in our human rights document, such as The UN Declaration.

clusion. When Arendt mentions the numbers of stateless persons in Europe after the first World War it became clear that their problem was not a matter of exception, on the contrary, it was simply a normal part of the way to organize the political life. Nation states exclude persons because on one hand they have the right to do so, they are sovereign. Moreover, in this specific case (imposed nation states, followed by an artificial creation of minorities), nation states also felt like they need to exclude. Arendt stresses that *Peace Treaties* ignored the necessary conditions demanded by a nation state, that is, homogeneity of population and its own territory. In her view, even when the necessary conditions are fulfilled, the establishment of a nation state presents its challenges. However, without it, universal human rights are doomed.

Another fact that Arendt brings to the conversation, is the practice of European nation states to denationalize unwanted persons. By doing that the nation states make sure that such persons were no longer their burden, given that they only have the obligation to protect their own nationals. Arendt highlights that the worse consequence of such practice was the fact that *everyone was* "*convinced* (…) that true freedom, true emancipation, and true popular sovereignty could be attained only with full national emancipation, that people without their own national government were deprived of human rights" (1962, 272). When Arendt says that everyone was convinced what she means is that the situation was accepted as normal, as simply as how things are. Given that, it was also accepted as normal the fact that some people will have no government to protect them and thus will be stateless, that is, rightless persons.

But Arendt's main argument is not that rightlessness became a byproduct of the nation state because people believe it to be a

normal phenomenon. People developed such an opinion because of the way that nation states reacted to the problem of statelessness in the first place. Given the big numbers of persons without a home or a government, the nation states' solution was to ignore the problem. By doing that, they make such persons superfluous. Later on, those ignored superfluous persons will be present to intermittent camps as the only house that the world has to offer them (cf. ibid., 279)

Arendt highlights that

> *the decision of the statesmen to solve the problem of statelessness by ignoring* it is further revealed by the lack of any reliable statistics on the subject. This much is known, however: while there are one million 'recognized' stateless, there are more than ten million so-called 'de facto' stateless; and whereas the relatively innocuous problem of the 'de jure' stateless occasionally comes up at international conferences, the core of statelessness, which is identical with the refugee question, is simply not mentioned. Worse still, the number of potentially stateless people is continually on the increase (Arendt 1962, 279; emphasis added).

If we apply Gündoğdu's reinterpretation of rightlessness to the problem of EVM today, then what Arendt just said in the quote above seems quite appropriate to explain the situation of EVM today. To a large degree, capable states are still ignoring their problem. We continue to live in a world where states have an almost absolute right to exclude. The only ones that we recognize as having a moral and a legal right to limit the state's right to exclude is the refugee in a juridical sense, who represents just a small part of the large group of superfluous persons here in cause. Moreover, in practice, even a smaller fraction will enjoy what I consider to be the moral rights of genuine refugees, that is, a chance to truly rebuild their lives with dignity, by having

the opportunity to belong to a political community, and consequentially having equality in front of the law and a chance to appear as *who* they are and not simply *what* they are (refugees or EVM).

In our days, the two most common solutions to deal with the protection of those that are legally recognized as refugees are to send them to camps in the Global South or to repatriate[77] them. In theory, the first solution is supposed to be temporary and to happen in dignified conditions, and the second is supposed to not impose forced repatriation and must take under consideration if it is save for the refugee to return. In practice, camps are very often the only home that the world has to offer to those superfluous persons. If we look at the facts, we will see that life conditions in refugee camps are often far from dignified. It is also true that the conditions under which repatriation happens are also in many circumstances in disaccord to the de-

[77] Those arguing for repatriation insist that we should not characterize refugees as stateless, because doing that constitutes a disservice to them, in the sense that undermines their right to return to their state (see Bradley 2014). Arendt, on the other hand, argues that to substitute the word stateless for the expression displaced persons is of interest for the nation states that decided to respond to such a problem by ignoring it. She argues that "Even the terminology applied to the stateless has deteriorated. The term "stateless" at least acknowledged the fact that these persons had lost the protection of their government and required international agreements for safeguarding their legal status. The post-war term "displaced persons" was invented during the war for the express purpose of liquidating statelessness once and for all by ignoring its existence. Nonrecognition of statelessness always means repatriation, i.e., deportation to a country of origin, which either refuses to recognize the prospective repatriate as a citizen, or, on the contrary, urgently wants him back for punishment" (Arendt 1962, 279).

mands of justice. Such facts, in my opinion, corroborate Arendt's and Shacknove's understandings of the problem of EVM. From a philosophical point of view, refugees are all the especially vulnerable persons whose most basic human rights are under threat and that because of that are in need of international protection. In my view, to take their plight seriously, philosophically, includes to acknowledge that their situation is fundamentally related to the fact that we live in a world where states have the right to exclude almost everyone. As a consequence of it, there are millions of superfluous persons forgotten beyond the margins of the world, in containment camps and detention centers. Forced to exist in such precarious conditions, they depend on charity to survive and are often exposed to arbitrarily decisions of the bureaucrats of states.

In my reading of Arendt, what is made explicit by her analysis of the phenomenon of the mass migration of a new type of refugee that appeared in Europe 80 years ago, is not that we must protect refugees because otherwise we would be putting in question the legitimacy of the state. Neither she justifies refugeehood by appealing to human's need or capacity to help. Her focus is on our right to matter, that is, on our right to not be made superfluous, or in other words, on our right to same moral worth. But she does not end the conversation by stating that all human beings have a right to matter or to be respected. She insists that in order to matter, one needs personhood. Arendt insists that the creation of superfluous persons is one unescapable aspect of organizing the world in states that have an almost absolute right to exclude. She insists that such byproduct of our statist politics, although so commonly among us, is morally wrong. On her own words to make human beings as human beings superfluous is the definition of evil. Moreover, in her view, all individuals have a moral duty to avoid evil. Or in other

words, her justification of why we should care about refugeehood in our days is because we have a moral duty to avoid evil. Next, I will discuss Arendt's account of evil. My main aim in this final part is to invite individuals to reflect about their own roles in addressing the problem of EVM.

3.

It is not so common to relate Arendt's remarks on refugeehood to her account of evil. Inspired by Patrick Hayden, in this paper I argue that Arendt's account of evil is very relevant today in order to help us understand problems such as extreme poverty and the problem of especially vulnerable migrants. Not only Arendt offers an explanation of why such problems persevere in our world, she also addresses the role of individuals in bringing change into the world. Arendt asks why evils like genocide and statelessness were possible in Europe 80 years ago. Her answer touches on a structural aspect of how we organize the world, and also on an ethical aspect of how human beings act.

The Moral Right in Arendt's Theory

Given that Arendt explicitly criticized morality and claimed that it had no use in politics, one could ask why bother looking for something that connects politics and morality in a theory that explicitly claims that these should not be related[78]. Arendt claimed that morality is a 'purity of heart' that has no room in politics (Arendt in Benhabib 1988, 46). Following Plato, she defined moral judgment as "the harmony or oneness of the soul with itself" (1971, 30). Arendt's view about morality is well exemplified in *On Revolution* by her remarks about the failure of the French Revolution, which in her view was a consequence of an incapacity to treat political problems as such. Poverty, for

[78] For Arendt's critique of morality, see Arendt (1990, 69-81); and Arendt's lecture "Thinking and Moral Considerations" (in 1971).

example, which should have been treated as a political problem, was approached as a matter of morality.

As Benhabib clarifies in this context, the morality to which Arendt directs her critique is the one that concerns itself with "the moral good", or in other words, with the "dispositions, traits of character, emotions, and intentions that lead to virtuous conduct" (1988, 46). Thus, one way of reading Arendt's critique of morality is as a critique of treating political problems as a matter of generosity and charity. According to Ayten Gündoğdu, "the lesson to be drawn from Arendt's account of 'the social question' is (…) that *a moralistic approach* to human suffering can easily turn into an instrumentalist understanding of politics as a means (if necessary, violent means) to achieve moral goals" (Gündoğdu 2015, 71)[79]. For Arendt, the French Revolution ended up in terror, because poverty, which should have been approached as a political problem (and therefore been communicate to people through solidarity), was made an object of 'the moral economy of compassion'[80]. That is, poverty became a problem to be solved by compassionate 'good hearts'

[79] Arendt's account of 'the social question', i.e. poverty, offered in her book *On Revolution* is controversial and it has been seen as a reason for accusing her of offering an 'humanitarian' account of human rights as the rights of helpless victims, unable to participate in politics and dependent on the help of others to survive. For such a critique see for example Rancière (2004). For a critique of Rancière's critique (see Stampone 2019). It is relevant to notice that if one considers only *On Revolution*, one may arrive at conclusions about Arendt's view on human rights, quite different than if one considers her whole oeuvre. Serena Parekh (2004) highlights that in a late essay called 'Public Rights and Private Interests' (1977), Arendt recognizes the fundamental importance of first fulfilling rights of subsistence in order for someone to be able to enjoy political rights.

[80] For a full account of this interpretation of Arendt's critique of the French Revolution, see Gündoğdu (2015, 67-75). Gündoğdu clarifies

and not as a matter of rights related to freedom and equality. According to Arendt, the consequence of this type of moralistic approach of politics is that matters of rights, end up being wrongly treated as a matter of charity or generosity[81].

What about the morality as a concern with what is morally right[82]? Does it have a place on Arendt's theory? I take the answer to be "yes". Although it is not something Arendt herself claimed, her theory seems to demand it. As Benhabib comments "[Arendt's] political theory of the public space, community, power, and participation seem to me to be inconceivable without *an implicit political ethics of enlarged thought*" (Benhabib 1988, 46). Despite Arendt's claim that morality and politics should be kept separate, she left an oeuvre that treats political problems as a matter of morality[83].

---

that she borrowed the expression 'moral economy' from Didier Fassin to whom the term captures "the production and circulation of a shared set of values and norms defining our moral world" (Fassin, "Compassion and Repression", 365 in Gündoğdu 2015, 226).

[81] In an Arendtian account, to approach poverty politically is to treat it as "an injustice incompatible with the foundations of freedom, one that relegates those who endure it into obscurity and denies them the right to political participation" (Gündoğdu 2015, 70).

[82] "The morally right concerns our public actions and interactions that affect, influence, and reflect upon the moral dignity and worth of the other as a public being" (Benhabib 1988, 46).

[83] By that I do not mean morality as a set of fixed values. In my view, for Arendt, morality is a matter of judgment. That is to say, politics are intrinsically related to the necessity of individuals been able to figure out what is right and what is wrong, that is, what is morally acceptable and what is not. For Arendt, the right way of doing it is by exercising the faculty of judgment, which is interrelated to the faculty of thinking (cf. Arendt 1971, 446).

Arendt's political thought is to a large extent motivated by what she experienced as a refugee and as a stateless person. One of her main concerns is what is necessary in order to avoid the evils of statelessness and genocide. She came up with the 'solution' that we need to demand 'the right to have rights' (cf. Arendt 1962, 296). Such 'solution' is in between invested commas not only because it is not a final one, what for Arendt is something dangerous and inefficient, but also because of Arendt's own pessimism about the feasibility of such paradoxical solution. As Parekh puts it, such 'solution' is more of a 'fruitful suggestion' (cf. Parekh 2004, 49). Contrarily to the tendency of our times, Arendt's defense of the right to have rights, is not linked to a defense of specific rights present in a list of Universal Human Rights. For example, if we would apply the Arendtian demand of the right to have rights to the problem of EVM, we would not end up with a list of human rights that could serve as justification of why states have duties of justice towards refugees.[84] In my view, Arendt acknowledges that states *should* have such obligations, in the sense that that is the right thing, the morally acceptable thing to do and to demand. What she does not acknowledge is the use of a theory that simply spits out such obligation. She refuses to do it in part because of her own experience as a refugee and as a stateless person. During these dark years she arrived at the conclusion that

[84] The right to membership in a political community is defended by Joseph Carens, who also believes in a human right to immigrate. Miller, on the other hand, accepts the human right to free movement, but not the human right to immigrate. They both acknowledge the human right to a decent standard of living as related to the refugees' rights. Moreover, the debate about the problem of EVM includes the reflection about whether or not many other rights should be considered human rights (cf. Carens 2013 and Miller 2016). In short, the discussion of the role of human rights in how we respond to the problem of EVM is a very important one, that I do not have the time to address here.

a political theory that simply draws ideal norms and values is useless and dangerous. Useless because to her, political actors need to think. Dangerous, because a set of 'good' values can easily be substituted by a set of 'evil' values, and those who are too used to obey, will simply keep obeying, without questioning it[85].

Because of her own experiences, Arendt decided that what was worth her time and her life was to engage in a type of political theory that is strongly realistic but also have some space for ideals. Its realism is a consequence of her decision that those thinking about political theory must see the world as it is, and not as they wish it to be (or were told that it was). That is why she insisted, for example, that human beings are not born equal (cf. 1962, 301) and that human rights were incapable of protecting human beings as human beings (deprived of their citizenship status) (cf. 1962, 293 f.). From that, it does not follow that all that her theory has to offer is despair. In Arendt's political theory one also finds ideals, which in my view, take the form of the 'fruitful suggestions' mentioned by Parekh. Although we are not born equals, we have the capacity to make ourselves equals through our decision as members of a group (cf. Arendt 1962, 301). In my opinion, the right to have rights is Arendt's highest example of such type of ideal. It is an ideal that does not have a well-defined form. It cannot be finally or fully explained by this or that human right. It can be interpreted in many different ways, and that is so, because after all, what Arendt does in her political theory, is to share perplexities. She admired Socrates, because of his capacity of helping others to think (cf.

[85] For Arendt that was to a large extent what happened in Germany during the Nazi regime (cf. Arendt 1971).

1971, 432 f.) and in my view, this is what she herself left us: an oeuvre that is an invitation for thinking[86].

Moreover, I believe, that such invitation is not an invitation to relativism. Arendt did believe that certain things are right, and by that, I mean morally acceptable, and others are wrong, that is, morally unacceptable[87]. What she did not believe was that the right way to approach politics and the world was by offering people a list of norms or values to follow. Human beings *need* to think and to judge, and by using these faculties, arrive by themselves at conclusions about what is right and what is wrong. That is why she ends up her discussion of the refugee question with the demand of the right to have rights. In my view, what Arendt does is to share perplexities, purposely and hoping, that this type of theory can serve as an invitation for thinking. One of the reasons why she invites human beings to think has to do with her believe that the only efficient way of avoiding *evil* necessarily goes through thinking. According to Arendt, it is of fundamental importance that individuals are able to think and judge and then by themselves figure out what is the right thing to do in a certain situation (cf. 1994, 321).

Drawing on Hayden, I assume that Arendt's notion of evil is useful to help us think about contemporary forms of superfluity. Hayden clarifies that from the fact that Arendt's "concept of evil derives from a specific focus on totalitarianism and genocide" does not follow that it does not have a relevance that "reaches beyond the atrocity of the Holocaust" (2010, 453). My

[86] For an account of Arendt's remarks on Socrates as a model of how to think see her essay "Socrates" (in 2005, 5-39).

[87] That becomes clear, for example, when she talks about those who chose to disobey the Nazi regime in order to do what was right.

claim is that the problem of EVM is a problem intrinsically related to what Arendt calls evil. By 'evil' Arendt does not mean something diabolic, neither something radical that only monsters would be able to commit. Evil is something that all humans are capable of, or as she calls it, evil is something banal (cf. 1965, 276). Because of that, it is of extreme importance that we do all that we can in order to avoid evil. She argues that what we can and should do is to exercise our capacities of judgment and thinking, which make it possible for us to *understand* the world we live in, and as such, to take into account all the human beings that are part of this world.

My claim here is that Arendt's notion of evil related to her notion of how we can understand a phenomenon that is happening in the present, gives us an account of ethics that does not tell us what to do, but instead tell us that we need to think about what is the right thing to do. Or in other words, in my opinion, although Arendt does not accept absolute moral truths in politics, her theory is inhabited by ethical concerns. Contrarily to most, Arendt's ethical concerns do not take the form of principles or values that guide our action. Instead, they are imbedded in her ideas of what evil is and of why we need to understand. Consequentially, what I am here referring to as an 'Arendtian ethics' is simply ethics in a broad sense, as a call for the need to think and judge in order to understand what right and wrong action is in front of a certain problem[88].

[88] I insist that from that does not follow any relativism. Arendt believed that the right thing to do is to act in a way that will not create rightless/superfluous persons. The right response to the Nazi totalitarian regime was to not be complicit. The ones that acted right were those who chose to disobey the Nazi government.

That said, my next task is to explain how Arendt understands our capacity to understand a phenomenon happening in the present. Arendt once said that fortunately, in order to be able to fight the phenomenon of totalitarianism, we don't have to understand it first (cf. 1994, 307). "Fortunately," because in her view, we are not capable of completely understanding a phenomenon while it is still happening. She insists that we are only able to understand it after it is over. In her words: "The understanding of political and historical matters, since they are so profoundly and fundamentally human, has something in common with the understanding of people: who somebody essentially *is*, we know only after he is dead (ibid., 309). But from that, it does not follow that we cannot understand the political problems that are part of our world in the present moment. According to her we can, and we *should* do so. She insists that if those fighting against totalitarianism want it to be more than a fight for survival, then they need to understand it (ibid., 310). In this sense, the type of ethics that I am defending is intrinsically related to our way of understanding. That is, we do understand phenomena happening in the present mainly to be able to decide what is the right action to take in front of it. Or in other words, in order to know what is the right thing to do in a certain situation we need to understand it. The next question then is how we can understand. Arendt insists that the tools that we must use to understand are simply our capacities to think and to judge and the facts of the world.

Arendt claims that with the rise of totalitarianism human beings lost their old tools of understanding (cf. ibid., 313). That is, they no longer can make use of "categories of thought and standards of judgment" such as cause and effect (ibid., 318). Totalitarian regimes showed to us that an 'event' that is something new appearing in the world is always unprecedent and unexampled (cf.

Vollrath 1977, 167). Because of that, we cannot truly understand a new phenomenon by applying to it our knowledge related to a previous phenomenon. Arendt mentions that that was what happened with people trying to understand totalitarianism by applying to this new phenomenon what they knew about imperialism (cf. 1994, 309-311). She calls what they were doing 'preliminary understanding'. For her understanding and knowledge are not the same but are intrinsically related. She explains that there is preliminary understanding and true understanding. The former precedes knowledge and the latter comes after it. Both types of understanding are what gives meaning to knowledge (cf. ibid., 310-11).

By applying Arendt's notion of understanding to the problem here investigated, I conclude that while we are still dealing with the problem of EVM here and now, all that we have in order to understand such a problem is our capacity to think and judge and the facts. By using such tools, we should preliminarily understand the problem and based on this preliminary understanding act accordingly. That is why I said in the beginning of this section that for Arendt to understand such a problem includes an ethical concern. Arendt insists that human beings are capable of acting in the right way by using their own capacity to judge informed by the facts of the world. She insists that "the actual fight against totalitarianism needs no more than a steady flow of *reliable information*" (ibid., 323 note 2)[89].

In my opinion, Arendt takes her criticism to normative categories too far, that is, I don't think that she is right to claim that they are never able to help us understand the phenomena that

[89] She continues by saying that "[i]f from these facts an appeal emerges, an appeal to Freedom and Justice, to mobilize people for the fight, then this appeal will not be a piece of abstract rhetoric." (ibid.).

happen in the world. However, I think that it is very important to listen carefully to her criticism, because it is also true that there are phenomena that cannot be correctly understood if we simply ask about its causes. What Arendt calls the phenomenon of statelessness and the condition of rightlessness is indeed one of these problems that cannot be explained through causality. What I have been calling the problem of EVM is also a problem of rightlessness, in the sense, that it is a political problem that reflects a systemic injustice that is consequent of the statist politics that we live in. Such a problem is in itself what Arendt considers to be evil, that is, "making human beings as human beings superfluous"[90] (Arendt and Jaspers 1992, 166). Because Arendt denounced evil in the context of her critique of totalitarianism, it is easy to misunderstand her notion of evil. As clarified by Hayden, Arendt's notion of evil is fundamentally related to her denunciation of "the logic of superfluity" that took over our statist politics after the First World War. For Arendt, such logic consists in "the modern expulsion [of rightless persons] from humanity" (Arendt 1948, 384). She explains that that is not only a matter of killing people, but first of all, a matter of dehumanizing them. As stressed by Hayden, "superfluous people are those cast out of a common world through the destruction of their political, legal, economic and moral status" (2010, 456).

In my view, Arendt's way of understanding, which as we saw, includes ethical concerns, is what we need in order to understand the specific problem here investigated. For her, the preliminary understanding of evil is intrinsically related to the

[90] To became superfluous for Arendt is to be in a fundamental condition of rightlessness. In this paper, I argue that although Arendt's original characterization of rightlessness is not applicable to help us understand the struggles of especially vulnerable migrants, a reinterpretation of it is.

fight to avoid evil. Those who think and judge and understand evil, will most likely choose to act in the right way. For her those who do not think will be mostly of the ones contributing to the maintenance of evil. That is in part why it is so important to do political philosophy on the problem of EVM in a way that asks what individuals can do to avoid such a problem, and not simply ask about what states are obliged to do[91].

Final Considerations

In my opinion, Arendt's accounts of refugeehood and evil are still quite relevant to help us understand and respond to the problem of EVM. Not only do they help us to better understand some important nuances of this political problem, they also remind us that we, as individuals, are also responsible for the protection of EVM. First, Arendt's account of refugeehood, focused in the problem as experienced by refugees, is a strong reminder that the expressions 'the refugee problem' and 'the problem of EVM' do not refer simply to problems that capable states are asked to solve. They refer also to the issues that refugees are forced to deal with, such as the loss of their homes, of their places in the world and of their distinctive features. Second, by bringing to the center of the discussion what does it mean to be forced to exist as a refugee, Arendt denounces that deep down the problem of EVM is a matter of making human beings superfluous. Third, she explains that the creation of large numbers of rightless persons is not a consequence of the failures of some states, but instead it is a structural problem that

[91] But from that, it does not follow that I consider the inquiry on what states owe to refugees and to other vulnerable migrants as a matter of justice irrelevant. I simply insist that we should embrace such inquiry carefully and taking under consideration that our primary role as political philosophers is to understand the problem. In my view, such a task in this specific case, is best achieved by doing an effort to understand in an Arendtian way.

is imbedded in our way of organizing the world (in states that have an almost absolute right to exclude). Fourth, she reminds us why it is important to care about refugeehood. To protect refugees is a matter of resisting a system that creates superfluous persons, or in other words, it is a matter of avoiding evil. At last, she insists that the fight against evil is constant, demanding, and it is an obligation of all of us. She explains that in order to fulfill our duty to avoid evil we need to imagine, to think, and to judge in each situation. We have a duty to consider the facts of the world and then use our mental capacities to figure out what is happening and how we must respond to the phenomenon that is taking place in our world. At last, I will reconstruct a partial sample of my thinking process regarding the response of European states to the refugee problem.

According to UNHCR it is estimated that there are 70.8 million forcibly displaced persons in the world today. Among those 25.9 million are refugees (in a juridical sense). The big majority of those forcibly displaced persons (90 per cent of them) live in developing countries (UNHCR 2019). To help us to have some perspective, European member states received 681.713 applications of Syrian refugees until the end of 2015. In the same period Turkey received 2.18 million applications of Syrian refugees (European Commission 2018, cf. Owen 2016, 142). With the exception of Germany, which is among the top hosting countries of refugees in the world today, most European countries are doing very little to protect refugees. That is to say, the facts show that European member states are not too generous towards refugees. On the contrary, they are failing to fulfill their obligation of justice towards them. While trying to justify such failure as something that is supposedly morally acceptable, they make use of the rhetoric that the fair thing to do is to share the burden of the refugee problem among all capable states. Although such idea is very present today in the mouths

of politicians and in the imaginary of European citizens, it is not present in the documents that constitute the refugee regime that we have in place. As clarified by Owen the current refugee regime "does not limit the duty to protect refugees to doing one's fair share of refugee protection" (2016, 143). What is established in those documents is that the protection of refugees should be efficient.

In order to illustrate that the refugee regime that we have in place prioritizes efficient protection of refugees over a fair distribution of the burden among capable states Owen discusses the principle of non-refoulement. He explains that "the duty of non-refoulement is a binding obligation on any state to which a claim to asylum is made not to return persons who, on the basis of an impartial process of adjudication, are found to satisfy the criteria of refugeehood to the state from which they have fled or to another state in which they would lack protection of their human rights" (ibid., 146). He highlights that one of the implications of this principle is that there is no established limitation of to how many refugees a state owes a duty of non-refoulement.

In my view, it must be worth coming up with a plan about how to share the burden of protecting refugees among the EU member states. Clear criteria about what constitutes the fair share of each state is necessary. But at the moment such criteria are not available. EU member states must be reminded that what is established on the 1951 Convention and on its 1967 Protocol is that capable states should prioritize protecting refugees. It is the job of individuals to remind states of it and to hold states accountable for their failures in protecting refugees and EVM.

Bibliography

Arendt, H. *The Origins of Totalitarianism*. New York: Meridian Books, 1962.

Arendt, H. *The Origins of Totalitarianism*. New York: Schocken, 1948.

Arendt, H. "'The Rights of Man': What are They?" *Modern Review* 3(1), 1949, 25-37.

Arendt, H. *Eichmann in Jerusalem*: *A Report on the Banality of Evil*. New York: Penguin, 1965.

Arendt, H. "Thinking and Moral Considerations". *Social Research* 38(3), 1971, 417–446.

Arendt, H. "Public Rights and Private Interests". *Small Comforts for Hard Times: Humanists on Public Policy*. M. Mooney and F. Stuber (eds.). New York: Columbia University Press, 1977.

Arendt, H. *On Revolution*. London: Penguin Books, 1990.

Arendt, H. "We Refugees". M. Robinson (ed.): *Altogether Elsewhere: Writers on Exile*. London: Faber and Faber, 1996, 111–119.

Arendt, H. *The Promise of Politics*. New York: Schocken Books, 2005.

Arendt, H. and Jaspers, K. L. "Letters". L. Kohler and H. Saner (eds).: *Hannah Arendt/Karl Jaspers: Correspondence 1926-1969*. New York: Harcourt Brace Jovanovich, 1992.

Benhabib, S. Judgment and the Moral Foundations of Politics in Arendt's Thought. *Political Theory* 16 (1), DOI 10.1177/0090591788016001003, 1988, 29–51.

Benhabib, S. *The Rights of Others*. Cambridge: Cambridge University Press, 2004.

Berstein, R. *Why Read Hannah Arendt Now*. Cambridge: Polity Press, 2018.

Bradley, M. "Rethinking Refugeehood: Statelessness, Repatriation and Refugee Agency" in *Review of International*

Studies 40, DOI:101017/50260210512000514, 2014, 101-123.

Brock, G. *Justice for People on the Move: Migration in Challenging Times*. Cambridge: Cambridge University Press, 2020.

European Commission. "Syria Crisis", http://ec.europa.eu/echo/files/aid/countries/factsheets/syria_en.pdf, 2018, accessed 22/01/2016.

Gündoğdu, A. *Rightlessness in an Age of Rights: Hannah Arendt and the Contemporary Struggles of Migrants*. Oxford: Oxford University Press, 2015.

Hayden, P. "The Relevance of Hannah Arendt's Reflections on Evil: Globalization and Rightlessness". *Human Rights Review* 11(4), DOI 10.1007/s12142-010-0157-8, 2010, 451–467.

Hayden, P. "From Exclusion to Containment: Arendt, Sovereign Power, and Statelessness". *Societies Without Borders* 3(2), DOI 101163/187219108X300046, 2008, 248–269.

Miller, D. *Strangers in Our Midst: The Political Philosophy of Immigration*. Cambridge, MA: Harvard University Press, 2016.

Owen, D. 2016. "Refugees, Fairness and Taking up the Slack". *Moral Philosophy and Politics*; 3 (2): 141-164.

Parekh, S. "A meaningful place in the world: Hannah Arendt on the nature of human rights". *Journal of Human Rights* 3(1), DOI: 10.1080/1475483042000185215, 2004, 41–52.

Rancière, J. "Who is the Subject of the Rights of Man". *The South Atlantic Quarterly* 103(2/3), DOI 10.1215/00382876-103-2-3-297, 2004, 297–310.

Saunders, N. "Hannah Arendt and the Refugee Problem". *International Political Theory and the Refugee Problem*. London: Routledge, 2017.

Shacknove, A.E. “Who Is a Refugee?”. *Ethics* 95(2), 1985, 274–284.

Shue, H. *Basic Rights*: *Subsistence, Affluence, and U.S. Foreign Policy*. Princeton: Princeton University Press, 1996.

Stampone, C. “Stateless, but Not Powerless: An Arendtian Perspective on Refugees and Power”. *Filosofisk Supplement,* September, 2019.

UN. “Declaration of Human Rights”, http://www.un.org/en/universal-declaration-human-rights/, accessed 09/06/2020.

UNHCR, https://www.unhcr.org/figures-at-a-glance.html, 2019, accessed 09/06/2020.

Vollrath, E. “Hannah Arendt and the Method of Political Thinking”. *Social Research* 44(1), 1977, 160–182.