ACTA PHILOSOPHICA FENNICA

Vol. 96

Normativity

The 2019 Entretiens of Institut International de Philosophie

Edited by
ILKKA NIINILUOTO
& SAMI PIHLSTRÖM

HELSINKI 2020

ISBN 978−951−9264−92−9
ISSN 0355-1792

Hansaprint Oy
Helsinki 2020

Table of Contents

Introduction

This volume of essays addressing philosophical issues of normativity from a variety of perspectives – including logical, epistemic, ethical, as well as historical – is based on the 2019 Entretiens of the Institut International de Philosophie (I.I.P.) organized at the University of Helsinki, Finland, on August 28 – 31, 2019.

Let us begin by warmly thanking everyone who contributed to making the I.I.P. conference possible. Obviously, without the formal decision by the I.I.P. itself, the event could not have taken place in Helsinki; accordingly, we are truly grateful to Professor Mircea Dumitru, the President of the I.I.P., and Professor Pascal Engel, the I.I.P. General Secretary, as well as the members of this Institute, for bringing the conference to Helsinki. The members of the Local Organizing Committee were the members of the I.I.P. based in Finland: Professors Lilli Alanen, Simo Knuuttila, Ilkka Niiniluoto, Gabriel Sandu (who is actually representing Romania as an I.I.P. member, though), and Sami Pihlström. In his role as the current President of the Philosophical Society of Finland, the latter acted as the Chair of the Local Organizing Committee. As organizers, we would like to warmly thank all these esteemed colleagues, as well as, of course, all the speakers – both members and non-members – who presented papers at the conference. The program was planned in collaboration by the I.I.P. leadership (President and General Secretary) and the local organizers.

Regarding the practical arrangements, such as reservations and budgeting, the contributions of Dr. Eero Kaila (Secretary of the Philosophical Society of Finland) and Ms. Päivi Seppälä (Treasurer of our Society) were absolutely fundamental; without their work the event simply would not have taken place. In addition, let us thank Mr. Panu-Matti Pöykkö, one of

our excellent doctoral students, for kindly acting as a conference assistant. Eero Kaila has also done marvellous work in copyediting this volume.

In addition to the I.I.P. itself, the conference had two local organizing institutions: the University of Helsinki and the Philosophical Society of Finland. We are deeply grateful to both for making the event possible. The conference team of the University and the Board of our Society crucially supported our arrangements; the University also made a substantial financial contribution to the conference by offering us the conference venue in Metsätalo and by hosting two receptions. Moreover, while the overall budget of the conference was modest, we did receive invaluable financial support from two sources: the Finnish Federation of Learned Societies and the Finnish Cultural Foundation. We would like to express our deepest gratitude for this crucial sponsorship.

The 2019 conference was the third time the official annual meeting of the I.I.P. took place in Helsinki during the prestigious history of the Institute beginning in 1937. The first I.I.P. conference in Helsinki was organized in 1970, with Academician Georg Henrik von Wright, one of the Past Presidents of the Institute, as the main local organizer; the topic was epistemology, and the conference volume is available as *Problems in the Theory of Knowledge*, ed. by G.H. von Wright (Martinus Nijhoff, 1972). 25 years later, in 1995, Ilkka Niiniluoto chaired the Local Organizing Committee, when the I.I.P. met to discuss "The History of Philosophy and the Methods of Philosophy"; the volume based on that conference was published with that same title, edited by Niiniluoto and Simo Knuuttila, in this very series, *Acta Philosophica Fennica*, in 1996. Another 24 years passed before our meeting in 2019 – though there was, during Jaakko Hintikka's presidency, also an I.I.P. meeting in Tartu, Estonia, with parts of the program arranged in Helsinki in 2001. The conference volume, edited by Matti Sintonen, was published with the title *The Socratic Tradition – Questioning as Philosophy and as Method* (College Publications, 2009). For more information about the I.I.P., its members, and its history since 1937, please see https://www.i-i-p.org/EN.

The fact that the I.I.P. once again organized its Entretiens in Helsinki can be regarded as a recognition of the international significance of Finnish philosophy. The local organizers

are obviously sincerely grateful to the I.I.P. and its members for this major acknowledgment, which is very highly appreciated by philosophers in Finland. Being a very small country, Finland has had several I.I.P. members, perhaps proportionally speaking even too many given the size of the philosophical community. However, all the Finnish members have been active in the I.I.P. in many ways and for a long time, starting with Eino Kaila in the early days of the Institute, followed by von Wright and Jaakko Hintikka (who was also a Past President), and later the already mentioned current members.

The Finnish tradition in philosophy is perhaps most widely known for its contributions to logic and the philosophy of science – merely names like von Wright, Hintikka, Niiniluoto, and Sandu speak for themselves. However, among the Finnish I.I.P. members, there are also leading historians of ancient, medieval, and early modern philosophy – Simo Knuuttila and Lilli Alanen. Clearly, the "Finnish tradition" is not restricted to any particular sub-field of philosophy. In Finland today, while foundational work in systematic analytic philosophy continues to be done at a high international level, there is also increasing heterogeneity in the field, with philosophers specializing in, say, phenomenology and pragmatism, as well as various historical topics.

Philosophy today is of course highly specialized – in Finland just as in most other countries – and some philosophers in the younger generations might not see much value in promoting general international collaboration in philosophy in the context of institutions such as the I.I.P. However, without analyzing this metaphilosophical issue at any length here, we wish to affirm our confidence in precisely this kind of general philosophical collaboration. In philosophy, the systematic as well as the historical issues in virtually all areas of research are almost always interrelated in a number of interesting ways, and while highly specialized research is of course needed in all the sub- and sub-sub- (etc.) fields of philosophy, there is something in the nature of philosophical activity itself that also motivates, or even requires, the pursuit of a general picture, a critical overview transcending the obvious and legitimate needs of specialization – both in terms of substance and in terms of the interplay of different national and region-

al traditions. Genuinely global institutions like the I.I.P. ought to be developed with such a metaphilosophical idea in mind, opening up further opportunities for leading philosophers to meet not only across national and geographical boundaries but also across the boundaries dividing philosophical traditions and specializations. This is not naïve promotion of interdisciplinarity (which is important in its own right but has arguably become a trendy catch-word in science policies inviting critical assessment) but on the contrary an affirmation of a certain kind of disciplinarity, an emphasis on philosophy as a discipline with its own intra-disciplinary boundaries that need to be critically explored and transgressed.

This brings us to the topic of our conference, *normativity*. It might be suggested that in the world we live in today, defending the normative sphere, or what Wilfrid Sellars called the "space of reasons", is a specifically philosophical (though of course also more broadly intellectual and cultural) task. We are all too familiar with attempts to reduce normative concepts and phenomena to mere contingent facts, or factual concepts. There is, as we all know, a tendency to dismantle rule- and norm-governed international institutions, practices, and agreements, replacing them with brute force and the power of the stronger. (Presidents Donald Trump's and Vladimir Putin's policies seem to be a clear manifestation of this tendency, but there are many other examples as well.)

The nine essays collected in this volume approach normativity from a variety of philosophical perspectives. (In addition to the articles published here, papers were delivered at the conference also by Professors Martin Kusch and Christoph Horn.) The volume opens with two historical papers exploring the medieval origins and developments of deontic logic (Simo Knuuttila) and Kant's critical philosophy, especially the Transcendental Dialectic of the First Critique, interpreted as "conceptual engineering" (Camilla Serck-Hanssen). A set of four papers is then devoted to broadly speaking logical and epistemic issues of normativity. Mircea Dumitru examines different philosophical and meta-logical ways of understanding logical normativity, while Wlodek Rabinowicz studies probability as the "value" of credibility based on a "fitting-attitude" analysis of value. Timothy Williamson then applies a non-modal framework of deontic normativity to the

norms of belief. Maria Lasonen-Aarnio, in turn, analyzes the problem of epistemic access to facts in virtue of which norms guide our action. The last group of three essays focuses on issues of (broadly speaking) practical normativity. Olav Gjelsvik criticizes attempts to reduce reasons to oughts, thus exploring the ontological foundations of (practical) normativity. Sara Heinämaa draws on von Wright's *Norm and Action* (1963) in seeking to clarify the conceptual confusions surrounding the vocabulary of "norms" and "normativity". Sami Pihlström's paper concludes the volume by asking a transcendental question about the very possibility of normativity as an irreducible dimension of human forms of life.

With these thoughts we hope that our readers will enjoy the articles based on the I.I.P. 2019 Entretiens.

Helsinki, July 2020

Ilkka Niiniluoto and Sami Pihlström

Historical Issues of Normativity

The Origins of Deontic Logic

SIMO KNUUTTILA

The most significant of the new theories put forward in early fourteenth-century logic were associated with modal logic and modal semantics, the previous approaches having been mostly dominated by Aristotle's views in Latin and Arabic logic. The new systematic thoughts were formulated in treatises on modal syllogistic, on consequences, on disputations logic that was called obligations logic in medieval times, and also in works on natural philosophy and theology. (See Knuuttila 2012; Lagerlund 2000; Thom 2003.) The intellectual attentiveness to modal theories and modal logic led authors to reconsider and modify Aristotle's modal syllogistic and other modal theories on the basis of the distinction between logical and natural modalities and other new insights into the meaning of modal notions, most importantly the idea of modality as alternativeness, which provided fourteenth-century theories with conceptual tools akin to some tenets of possible-worlds semantics and related modern interpretations of modality. These elaborations have been studied by many historians of logic since the 1980s. While late medieval modal theory is an interesting part of the history of logic and also congenial to the style of analytic philosophy, I shall not address its main features in any more detail, but instead attend to what could be called applied modal logic, to use Ockham's suggestion (*Summa logicae* II.1, 243.67-68), particularly the beginnings of deontic logic. The modern development of this branch of logic was initiated by G.H. von Wright's "Deontic Logic" (1951) and it has been continued by many Finnish and Scandinavian philosophers and logicians. (See the two volumes on deontic logic edited by R. Hilpinen: 1981, 1981a.)

1. Medieval applied modal logic

In explicating their new guidelines on modal logic proper, nowadays called alethic modal logic, late medieval thinkers realized that there were philosophically important notions, such as knowledge and belief, the logical behavior of which was analogous to that of the basic modal notions of necessity and possibility. This made them think and write about the similarities between the logic of alethic and epistemic modalities in their works on modal logic and separate treatises on epistemic terms. Fourteenth-century epistemic logic came to be developed somewhat differently by the English logicians such as William Ockham and William of Heytesbury and the Parisian John Buridan and their followers (Boh 1993; Knuuttila 2015). An analogous approach to deontic notions of obligation and permission was also suggested in this context; while this was not as popular as epistemic logic, some visible authors such as Robert Holcot and Gregory of Rimini discussed the principles of deontic consequences. (Robert Holcot, *Quodlibet I*, f. 152ra-152vb; Gregory of Rimini, *Lectura super primum et secundum Sententiarum*, II.38-41, pp. 304-305). A special treatise on this subject was written by Roger Roseth in the 1330s; he was an English author influenced by the English discussions of logic after William Ockham. There is no modern edition of Roseth's treatise on deontic concepts, but a longer part of it was printed under the name of Holcot in the opening question of *Determinationes magistri Roberti Holcot* (Lyons 1518). I have translated with Olli Hallamaa parts of Roseth's chapter about deontic matters in *Lectura super Sententias* (q. 1, a. 2) from an unedited manuscript (see Knuuttila & Hallamaa 1995).[1]

Medieval modal logic and its epistemic and deontic extensions were partially known to some sixteenth- and seventeenth-century authors, but these theories were not furthered because logic had lost its central status in philosophy, one of the few exceptions being Leibniz who made some comments on medieval logic and also on deontic concepts (Lenzen 2004, pp. 320-324). Therefore, the title of this paper, which alludes

[1] Some other parts of Roseth's commentary on the *Sentences* are edited in Hallamaa 2005. In what follows, the references to Roseth's deontic logic are mostly to the translation in Knuuttila & Hallamaa 1995.

to medieval origins of applied modal logic, does not refer to the beginning of a continuous tradition. It points to historical investigations of logical themes that were intended to deal with questions and problems largely analogous to those in the treatises on epistemic logic and deontic logic introduced in the 1950s. It is not very surprising to find these accounts in the fourteenth century – there were at that time scholars coming from a scholarly tradition of many generations trained in logic at medieval universities. Having a powerful new modal theory, they were motivated to apply their logical insights in new areas. It seems that they came to applied modal theories by attending to conceptual analogies in a way similar to what was described in an autobiographical note by G.H. von Wright (2001, pp. 171-174).

A central topic in medieval epistemic logic was the question about the epistemic closure under deduction

(1) $K\varphi \,\&\, (\varphi \rightarrow \psi) \rightarrow K\psi$.

A related principle was the epistemic closure under known deduction

(2) $K\varphi \,\&\, K(\varphi \rightarrow \psi) \rightarrow K\psi$.

These two forms have been often compared in contemporary epistemic logic since Jaakko Hintikka's book *Knowledge and Belief* (1962), where (1) is defended for conceptual and systematic reasons. Medieval authors preferred (2) to (1), and this medieval restriction was approvingly stressed by R.M. Chisholm in his review of Hintikka's book (Chisholm 1963b, pp. 773-795). Things were not quite that simple because in the medieval theories of the consistency of questions and answers, developed in obligations logic, the answers were evaluated on the basis of (1), without an explicated knowledge of the consequence; the acceptance of logical consequences could be taken for granted in logical discussions. This view was also found in Aristotle's *Prior analytics* II.21.[2]

In medieval logic, modalities were divided into compounded and divided ones (*in sensu composito*, *in sensu diviso*), sometimes also called modalities *de dicto* and *de re*. In the compounded sense the modal notion modalized the significa-

[2] Knuuttila 2019, pp. 85-102.

tion of a modal proposition and in the divided sense it modalized the copula, expressing how the subject was what it was thought to be. Similarly, a compounded knowledge proposition meant that the truth of an understood proposition was known and the divided proposition that about the subject of the proposition it was known what was predicated of it, the subject being outside the scope of knowledge. Compounded knowledge propositions were typically read internally, from the point of view of the knowing subject, and divided knowledge propositions were read externally.[3] This division played an important role in medieval epistemic logic, particularly in distinguishing between "knowing who" and "knowing that" and related terms in singular propositions with a demonstrative pronoun. In spite of the popularity of this analysis of the structure of modal propositions, it was not applied to the logic of deontic propositions. I shall now turn to this theory.

2. Defining deontic notions

The following equivalences analogous to those between the notions of possibility, necessity, and impossibility were used in the fourteenth-century discussions of normative terms:

---

[3] Knuuttila 2008, pp. 533-536, 551-559. According to William Heytesbury, epistemic modals in the compounded sense did not imply those in the divided sense or *vice versa*, with the exception of singular epistemic predications with respect to the demonstrative pronoun "this" used in a definite way; see Heytesbury's *The Verbs 'Know' and 'Doubt'*, translated in *The Cambridge Translations of Medieval Philosophical Texts I*, ed. N. Kretzmann & E. Stump 1988, pp. 443-446, 454-455. See also his *The Compounded and Divided Sense,* ibid., pp. 426-432. In fourteenth-century modal logic, the compounded and divided senses were taken to be equivalent in this particular case of propositions with a demonstrative pronoun as a subject. See William Ockham, *Summa logicae*, II.10 (pp. 276-279); III-1, 32 (p. 448); III-3, 10 (pp. 632-634); John Buridan, *Tractatus de consequentiis*, II.7, 16 (pp. 75-76). For Heytesbury's principles in Paul of Venice's *Logica magna I, Tractatus de scire et dubitare*, ed. and trans. by P. Clarke (*On knowing and Being Uncertain*), 1981, pp. 52-53, 116, 142-148. The same questions are often addressed in Peter of Mantua's *Logica* (Padua 1477); he discussed epistemic notions by taking Heytesbury's treatise as a starting point. See also Strobino & Knuuttila 2020.

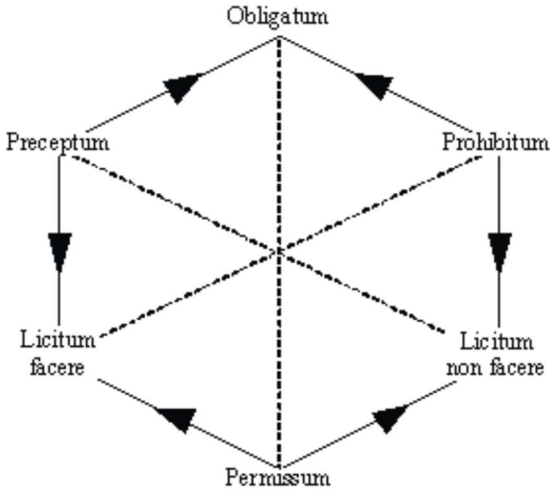
(3) Pp = -O-p = -Fp

(4) P-p = -Op = -F-p

(5) Op = -P-p = F-p

(6) O-p = -Pp = Fp.

(O stands for obligation (*obligatum*), P for permission (*licitum*) and F for prohibition (*illicitum*).) Most of these forms or their parts are found, for example, in Ockham's *Opus nonaginta dierum*.[4] While many of these equivalences were employed in legal texts and treatises on natural laws and rights, they were seldom discussed separately and not included in the popular square of modal opposition.[5] However, there were some systematic studies of their logical relations, most explicitly in Marsilius of Padua's *Defensor pacis* from early fourteenth century.[6] In this work they are arranged in a way that can be presented as follows (II.12, pp. 3-4):



This deontic hexagon, put together by Brian Tierney from Marsilius's formulations, is also found in Gottfried

[4] William of Ockham, *Opus nonaginta dierum*, in *Opera politica* I-I, 1940, 1963 For obligation and permission and right as a licit power in medieval natural rights theories, see Kilcullen 2010; Tierney 2014.

[5] For the square of modal oppositions, see Knuuttila 2008, pp. 531-533, 554.

[6] Marsilius of Padua, *Defensor pacis*, ed. R. Scholz, 1932-1933.

Achenwall, an 18th-century German author who influenced Kant's legal philosophy.[7] According to Marsilius, if one has a positive obligation (*obligatum*) to do something (*praeceptum*), one is permitted to do it (*licitum facere*), and if one has a negative obligation concerning something (*prohibitum*), it is permitted not to do it (*licitum non facere*). As distinct from alethic possibility and necessity, actuality did not imply permission and obligation did not imply actuality. These forms of being obliged or permitted were also used in Roger Roseth's treatise, as well as the consistency principle that if one has an obligation to do something according to a system of norms, one cannot simultaneously have an obligation to omit it. Deontic consistency is defined by Roseth as the basis of the rationality of norms.[8]

Because of the logical properties of deontic notions, it was natural to ask about the validity of the deontic closure for obligation and permission:

(7) $O\varphi \,\&\, (\varphi \rightarrow \psi) \rightarrow O\psi$.

(8) $P\varphi \,\&\, (\varphi \rightarrow \psi) \rightarrow P\psi$.

These forms, similar to those of epistemic closure, derived from the extensive fourteenth-century discussions of consequences and inferences. As mentioned above, (7) and (8) were addressed by Robert Holcot and Gregory of Rimini and, in a more extensive way, by Roger Roseth. None of these authors accepted (7) or (8) without qualifications.

3. Deontic closure

Late medieval theory of deontic logic was preceded by a discussion of the logic of will since the twelfth century. This logic was treated propositionally, "willing that p" meaning that an agent has an informed and free intention to do what is signified by the proposition. It was asked whether a principle similar to (8) can be applied to the notion of will. The tradition of the logic of will was relevant to Roseth who developed his deontic theory for volitions from the point of view of divine law.

---

[7] See Tierney 2014, p. 134.

[8] For the rules of rationality, Knuuttila & Hallamaa 1995, pp. 79-83.

Some authors wrote that when the consequent of an inference expresses a necessary means to the antecedent, willing the antecedent implies willing the consequent.[9] This is what Immanuel Kant characterized as an analytical truth of hypothetical imperatives: if one does not will the consequent, one does not will the antecedent.[10] This use of the notion of will was similar to the deontic closure in the sense that willing the end implies willing a means that is necessarily required for the end. However, medieval authors noticed that the consequent of what was willed was not necessarily a means, but could also be something else that was implied by the antecedent and not willed when the antecedent was willed.

Peter of Poitiers (c. 1170) wrote that while some of his contemporaries held that knowing the antecedent of a good consequence implied knowing the consequent, they denied the corresponding principle about the notion of will, stating that "If something follows from something, it does not follow, as with knowing, that he who wills or can do the antecedent wills or can do the consequent. They want to show this with logical and theological examples ... If this person is in some way in Rome, he is in Rome, and he wills to be in some way in Rome, but he does not will to be in Rome". (*Sententiae I*, ed. P.S. Moore and M. Dulong 1961, I.9, pp. 82-83.) The author does not explain what kind of will to be in Rome does not imply a will to be in Rome, but an analogous example was later often treated in discussing ambiguous propositions. The question of willing the antecedent and the consequent was then associated with the consequence: if one is in the mud with 100 pounds, one is in the mud. William Ockham and Robert Holcot took this to refer to a case of being paid for doing something in the mud, in which case it was reluctantly willed but willed any way as a rewarded act.[11] Second, one might will the antecedent and not directly intend to stick in the mud, although it might be a known but unintended side-

---

[9] See William Ockham, *Quaestiones variae, ed. G. Etzkorn et al.*, 6.9, p. 259.

[10] *Grundlegung zur Metaphysik der Sitten*, in *Werke*, Vol. VI, ed. W. Weischedel, p. 46.

[11] See Ockham, *Summa logicae* III-4, 13 (p. 836); Holcot, *Exploring the Boundaries of Reason* 84.525-86.563.

effect of action, for example, will to grab one's purse out of the mud. Without using this example, the case was particularly analyzed by Thomas Aquinas in his discussion of self-defense (*Summa theologiae* II-2, 64.7), which gave rise to the influential doctrine of double effect.[12] The third alternative could be that one is unwillingly in the mud and then finds 100 pounds there, thinking that he wills to be in the mud with 100 pounds rather than without it. In this case he did not will to be in the mud at all.[13]

The last kind of possibility was mentioned by Peter of Poitiers in returning to his willing to be in Rome example and combining it with the doctrine of repenting of one's sin, which was later similarly employed by Roger Roseth.[14] Roseth writes: "This consequence is good and known to be good: I repent of my sin; therefore I am guilty of sin. I am permitted to will the antecedent, but I am not permitted to will the consequent, because I am permitted to repent of my sins, but I am not permitted to will to be guilty of sin. Therefore this consequence is not valid."[15] Roseth understood formal consequences in the same way as many of his contemporary Englishmen: what is signified by the antecedent entails what is signified by the consequent.[16] When it is understood that Socrates is repenting of a sin, he is understood to be guilty of sin. In Roseth's view, the deontic consequence principle is not applicable here because repentance is required for fulfilling one's remedial obligation that is actualized by a previous violation of a duty. According to Roseth, avoiding sins is a duty and repenting them is a dependent duty. This analysis is similar to what is found in the discussion of contrary-to-duty imperatives in modern deontic logic since Chisholm's paper from 1963 that restricted the validity

---

[12] See also Matthews 1998. According to William of Auxerre (d. 1231), being in the mud was not willed in the divided sense of the example (*Summa aurea* III.17, pp. 2, 5).

[13] This may be meant by Walter Burley who simply denies that willing the antecedent implies willing the consequent; see *De puritate artis logicae* 206.34-207.2. See also Knuuttila & Holopainen 1993, pp. 115-132.

[14] Peter of Poitiers, *Sententiae* IV, 16, *Patrologia Latina*, ed. J.P. Migne, vol. 211, p. 1199.

[15] See Knuuttila & Hallamaa 1995, p. 84.

[16] See Read 2010, pp. 173-184.

of standard deontic logic with deontic closure. Contrary-to-duty imperatives tell what people ought to do if they have violated their duties; they are obligations that come into effect when one has acted contrary to primary obligations.[17]

Another twelfth-century example for refuting the general validity of the consequence principle for the notion of will was formulated by Stephen Langton (c. 1200) as follows: Necessarily, if a man visits his sick father, the father is sick. But it does not follow that if a man wills to visit his sick father, he wills the father to be sick. "He wills what implies, namely visiting his sick father, but he does not will what is implied". This is said to refute the principle that "if a consecutive antecedent is good, namely that from which something follows, the consequent is good".[18] One ought to will to visit the sick father because it is a good thing, but this does not imply that one ought to will that the father is sick. This is in fact forbidden. Langton's example is similar to what is called the paradox of the Good Samaritan in recent discussions.[19] Roseth does not discuss the sick father case, but he addresses some others with the same idea; for example, when sleeping one does not do any good things and it is often permitted to will to sleep but not permitted to will that one does not do any good things. Roseth's other examples based on the distinction between what is permitted and permitted to be willed and what is permitted but not permitted to be willed include the cases of being permitted to be a son of a criminal father, but not permitted to will this, and being permitted to be without money, which implies that one cannot give alms, but it is not permitted to will not to give alms. Roseth's deontic rules basically apply to voluntary acts and omissions, but the catego-

[17] Chisholm 1963a. See also von Wright 1991, pp. 105-120.

[18] See the text in Quinto 1992, pp. 127, 129-130.

[19] See also Hilpinen 2019, pp. 420-434. Hilpinen wonders why medieval authors did not apply their distinction between the composite and divided readings of modal propositions to determine the deontic scope. He argues that this is the shortcoming in their counter-examples to the deontic consequence principle as well as in some contemporary deontic paradoxes.

ry of things permitted but not permitted to be willed implies a distinction between permitted to be and permitted to do.[20]

The so-called paradoxes of contrary-to-duty imperatives and the good Samaritan are among the most quoted problems concerning the deontic consequence principle in modern studies, but interestingly the same type of examples were employed by Roger Roseth in evaluating the validity of deontic inferences. It was already stated by Peter of Poitiers that an important aspect of the logic of will is that, as distinct from epistemic logic, the consequence principle does not hold of the notion of the will. According to Roseth, it is only applicable with specific restrictions in deontic logic, the deviations being just these two types of imperatives. They apparently codify the most natural problematic cases in considering the inference in deontic modal logic.

Because of his criticism, Roseth was motivated to consider the concept of conditional obligation. His preliminary idea was that q may be obligatory on the condition that p is actual:

(9) $p \rightarrow Oq$,

but he argued in a pretty extensive paragraph that something more is needed here. He first deals with a sophistical argument, in which it is assumed that p and q signify the same and the obligation is fulfilled when the condition is fulfilled. As part of this discussion he introduces a conditional norm that there is an obligation to repent if and only if is one is guilty. If Socrates repents when he is not guilty, he violates the norm and apparently ought to repent, but this is what he is doing and hence apparently ought not to repent. So if he ought not to repent, he ought to repent and vice versa. To avoid difficulties of this kind, Roseth suggests that one should add to the conditional form the qualification that the actuality of p is not *eo ipso* accompanied by the actuality of q:

(10) $p \rightarrow Oq \ \& \ -(p \rightarrow q)$.[21]

---

[20] See Knuuttila & Hallamaa 1995, pp. 80, 83-84. Cf. also the contemporary discussion of ought-to-be and ought-to-do in d-Altan, Mayer, Wieringa 1996, pp. 77-111.

[21] Knuuttila & Hallamaa 1995, p. 85; for the same condition as modalized - $N(p \rightarrow q)$, see von Wright 1981, p. 169.

Furthermore, he would like to add the condition that the conjunction of p and q is possible. So the final form is as follows:

(11) $p \rightarrow Oq \,\&\, \neg(p \rightarrow q) \,\&\, M(p \,\&\, q)$.

His example of the last part is that Socrates ought to cross a bridge, if he says something true. Socrates says that I shall not cross the bridge, which if true implies that the obligation cannot be fulfilled. This remark was associated with a longer discussion of *insolubilia* which was one of the favorite topics of fourteenth-century logicians.[22]

University of Helsinki

References

d'Altan, P., J.-J. Mayer, & R.J. Wieringa (1996), "An Integrated Framework for Ought-to-Be and Ought-to-Do Constrains", *Artificial Intelligence and Law* 4, pp. 77-111.

Boh, I. (1993), *Epistemic Logic in the Later Middle Ages*, Routledge, London and New York.

Chisholm, R.M. (1963), "Contrary-to-duty Imperatives and Deontic Logic", *Analysis* 24, pp. 33-36.

Chisholm, R.M. (1963), "The Logic of Knowing", *Journal of Philosophy* 60, pp. 773-795.

Gregory of Rimini (1980), *Lectura super primum et secundum Sententiarum*, ed. V. Marcolino, W. Simon and V. Wendland, vol. VI, de Gruyter, Berlin.

Hallamaa, O. (2005), *Lectura super Sententias: Quaestiones 3, 4, & 5*, Luther-Agricola-Seura, Helsinki.

Hilpinen, R. (ed., 1981), *Deontic Logic: Introductory and Systematic Readings*, Reidel, Dordrecht.

Hilpinen, R. (ed., 1981a), *New Studies in Deontic Logic*, Reidel, Dordrecht.

Hilpinen, R. (2019), "On the Sick Father, The Repentant Sinner, and Other Problems in Medieval Deontic Logic", *Theoria* 85, pp. 420-434.

Hintikka, J. (1962), *Knowledge and Belief: An Introduction to the Logic of Two Notions*, Cornell University Press, Ithaca, NY.

John Buridan (1976), *Tractatus de consequentiis*, ed. H. Hubien, Publications Universitaires, Louvain; Vander-Oyez, Paris.

[22] See Roger Roseth, *Lectura super Sententias*, Oxford Oriel College 15, 252ra; for *insolubilia*, see also Rahman, Tulenheimo, and Genot 2008.

Kant, I. (1983 [1785]), *Gundlegung zur Metaphysik der Sitten*, in *Werke*, Vol. VI, ed. W. Weischedel, Wissenschaftliche Buckgesellschaft, Darmstadt.

Kilcullen, J. (2010), "Medieval and Modern Concepts of Rights: How Do They Differ?" In: V. Mäkinen (ed.), *The Nature of Rights: Moral and Political Aspects of Rights in Late Medieval and Early Modern Philosophy*, Acta Philosophica Fennica 87, Societas Philosophica Fennica, Helsinki, pp. 31-62.

Knuuttila, S. (2008), "Medieval Modal Theories and Modal Logic", in *Handbook of the History of Logic II: Mediaeval and Renaissance Logic*, Elsevier, Amsterdam, pp. 505-578.

Knuuttila, S. (2012), "Modality", in J. Marenbon (ed.), *The Oxford Handbook of Medieval Philosophy*, Oxford University Press, New York, pp. 312-341.

Knuuttila, S. (2015), "Epistemic Logic in Paul of Venice's Commentary on the Beginning of *Posterior Analytics*", in. J. Biard (ed.), *Raison et démonstration. Les commentaires médiévaux sur les Seconds Analytiques*, Brepols, Turnhout, pp. 185-198.

Knuuttila , S. (2019), "Epistemic Logic in Aristotle", in J.I. Lindén (ed.), *Aristotle on Logic and Nature*, Peeters, Leuven, pp. 85-102

Knuuttila, S. & O. Hallamaa (1995), "Roger Roseth and Medieval Deontic Logic", *Logique et analyse* 38, pp. 75-87.

Knuuttila, S. & T. Holopainen (1993), "Conditional Will and Conditional Norms in Medieval Thought", *Synthese* 96, pp. 115-132.

Lagerlund, H. (2000), *Modal Syllogistics in the Middle Ages*, Brill, Leiden.

Lenzen, W. (2004), *Calculus universalis. Studien zur Logik von G.W. Leibniz*, Mentis, Paderborn.

Matthews, G.B. (1998), "Saint Thomas and the Principle of Double Effect" in S. MacDonald and E. Stump (eds.), *Aquinas's Moral Theory*, Cornell University Press, Ithaca, NY, pp. 63-78.

Marsilius of Padua (1932-1933), *Defensor pacis*, ed. R. Scholz, Hahn, Hannover.

Paul of Venice (1981), *Logica magna I, Tractatus de scire et dubitare*, edited with an English translation in *On knowing and Being Uncertain* by P. Clarke , The British Academy/Oxford University Press, Oxford.

Peter of Mantua (1477), *Logica*, Padua.

Peter of Poitiers (1961), *Sententiae I*, ed. P.S. Moore and M. Dulong, University of Notre Dame Press, Notre Dame.

Peter of Poitiers, *Sententiae*, *Patrologia Latina*, ed. J.P. Migne, vol. 211.

Rahman, S., T. Tulenheimo & E. Genot (eds.) (2008), *Unity, Truth and the Liar: The Modern Relevance of Medieval Solutions to the Liar Paradox*, Springer, Dordrecht.

Read, S. (2010), "Inferences". In R. Pasnau and C. van Dyke (eds.), *The Cambridge History of Medieval Philosophy I*, Cambridge University Press, Cambridge, pp. 173-184.

Robert Holcot (1518), *Determinationes*, Lyons.

Robert Holcot (1983), *Exploring the Boudaries of Reason: Three Questions on the Nature of God by Robert Holcot*, ed. H.G. Gelber, Pontifical Institut of Mediaeval Studies, Toronto.

Robert Holcot, *Quodlibet I*, Ms. London, British Library, Royal 10 C VI, f. 141vb-173vb.

Roger Roseth (2005), *Lectura super Sententias*, ed. O. Hallamaa, Luther-Agricola-Seura, Helsinki.

Roger Roseth, *Lectura super sententias*, Ms. Oxford, Oriel College 15, 235r-279v.

Strobino, R. & S. Knuuttila (2020), "William Heytesbury and Peter of Mantua on Demonstrative Pronouns in Epistemic Contexts", forthcoming.

Quinto, R. (1992), "*Die Quaestiones* des Stephan Langton über die Gottesfurcht", *Cahiers de l'Institut du Moyen-Age grec et latin, Université de Copenhague* 62, pp. 77-165.

Thom, P. (2002), *Medieval Modal Systems: Problems and Concepts*, Ashgate, Aldershot.

Tierney, B. (2014), *Liberty and Law*, The Catholic University of America Press, Washington, DC.

von Wright, G.H. (1951), "Deontic Logic", *Mind* 60, pp. 1-15.

von Wright, G.H. (1981), "A New System of Deontic Logic", in Hilpinen (ed.), *Deontic Logic*, pp. 105-120.

von Wright, G.H. (1981), "Deontic Logic and the Theory of Conditionals", in Hilpinen (ed.), *Deontic Logic*, pp. 159-177.

von Wright, G.H. (2001), *Mitt liv som jag minns det*, Bonniers, Stockholm.

William of Auxerre (1980-1987), *Summa Aurea*, ed. J. Ribaillier, CNRS, Paris.

William Heytesbury (1988), *The Compounded and Divided Sense* and *The Verbs 'Know' and 'Doubt'*, translated in *The Cambridge Translations of Medieval Philosophical Texts. Volume one: Logic and the Philosophy of Language*, ed. N. Kretzmann and E. Stump, Cambridge University Press, Cambridge, pp. 415-434, pp. 436-479.

William Ockham (1974), *Summa logicae*, ed. Ph. Boehner, G. Gàl and S. Brown, St Bonaventure University, St Bonaventure, NY.

William Ockham (1984), *Quaestiones variae, ed. G. Etzkorn et al.*, St Bonaventure University, St Bonaventure.

William of Ockham (1940-1963), *Opus nonaginta dierum*, *Opera politica*, vol. I, ed. R.F. Bennet and J.G. Sikes, vol. II, ed. R.F. Bennet and H.S. Offler, Manchester University Press, Manchester.

Kantian Critique as Conceptual Engineering

CAMILLA SERCK-HANSSEN

Introduction[1]

This paper is divided into three main parts. In the first part I will present the main points of a new philosophical field called conceptual engineering. An essential feature of conceptual engineering is its emphasis on philosophy as a normative discipline. In part two I will explain why I take Kant to be a conceptual engineer and suggest how his position can be placed *vis a vis* the contemporary debate in that field. In part three I will argue that Kant engineers the concept of <negation> and that his critique of speculative metaphysics in the Transcendental Dialectic of the *Critique of Pure Reason* (CPR) can be shown to rest on this engineering. To prove my case, I will show in some detail how my reading works with Kant's critique of two of the paralogistic inferences through which the speculative metaphysician attempts to deduce that the soul is a simple substance.[2]

What is conceptual engineering?

Less than a decade ago a new field of philosophy was put on the agenda: conceptual engineering. Roughly speaking, conceptual engineering aims at criticizing and improving concepts. It involves *diagnosing* shortcomings such as ambiguity, vagueness, superimposition, inconsistency and emptiness (Cappelen 2018; Plunkett and Cappelen forthcoming; Eklund 2017) as well as shortcomings of the social and political kind.

---

[1] I thank an anonymous reviewer of an earlier version of this paper for insightful comments and questions concerning Kant's normative stance.

[2] This part of the paper draws on material published in Serck-Hanssen (2019) and Serck-Hanssen (2017).

The latter are argued to arise through concepts which preserve oppression of gender and social inequality (Haslanger 2012), or concepts that foster unjustified generalizations and prejudice (Leslie 2017; Sterken 2019). Conceptual engineering aims at ameliorating and sometimes at replacing concepts found to be wanting or defective. Improvement may take place through different strategies such as disambiguation, splitting and revision of conceptual content.

Conceptual engineering is the offspring of philosophy of language. It is a reaction to the last 50 years of work within that field and a critique of its thoroughgoing *descriptive* method.[3] Rather than the good old attempt at describing language and linguistic behavior, conceptual engineering is a *normative* endeavor. It aims at fixing language, at improving our representational devices or concepts (Cappelen 2018). Ardent conceptual engineers argue that to engage in this kind of intellectual activity is what philosophers not only can, but indeed also should, do (Eklund 2014, p. 293).

Within the debates about conceptual engineering there are three issues of particular controversy. The first concerns the object of engineering. One might assume that the answer is obvious: what conceptual engineering engineers is of course a concept! But there is in fact a lively discussion about whether this is the correct view. Some of the most influential voices in the field maintain that conceptual engineering should not be concerned with concepts or the conditions for concept possession. It is instead an engineering of the object level or the world (Cappelen 2018). Other conceptual engineers argue that what is engineered are neither concepts nor objects, but rather words and their meaning (Plunkett 2015). But there are also conceptual engineers such as Eklund (2014) and Scharp (2013) who take conceptual engineering to be directed at concepts or more precisely their constitutive principles. On their view, the main aim of conceptual engineering is to fix incon-

[3] Some might object to this way of describing the field. There are strands within philosophy of language that lean on a certain reading of Wittgenstein's philosophy, which appears to have more in common with the normative view advocated by conceptual engineering. In this paper I will, however, not enter into this debate and simply take the standard description of the field, as offered by e.g. Cappelen, for granted.

sistent concepts, i.e. concepts whose constitutive principles jointly entail falsehoods.

A second issue concerns the very idea of engineering. Some people argue that whatever the object is that is "worked" at, the work cannot be conceived of as something akin to engineering, for engineering has a clear "recipe" for *how* things are to be put together so as to yield the required result, say a car engine or a bridge or a computer. But none of the conceptual engineers have come up with a method for their activity. Second, the success of ordinary engineering is basically only dependent upon the skills of the engineer (and on having the right material or parts at hand) but since conceptual change dovetails with linguistic change and linguistic meaning is determined externally, the idea that one person could succeed as a conceptual engineer is utopian.[4]

These two challenges can, however, be met in the following way. First, while it is correct to say that no one has come up with a recipe for the engineering of concepts, this does not show that the philosophical field referred to by "conceptual engineering" is misguided. It might be true that the label is not fully adequate and can give false expectations, but not much hangs on that choice. Some people who conceive of themselves as dedicated to this field of philosophy prefer to call what they are involved in "conceptual ethics" (Burgess and Plunkett 2013a, 2013b). In other words, one should not take the word "engineering" to suggest that conceptual engineers need to be committed to conceptual engineering being similar to ordinary engineering in the way the objection presupposes. If one succeeds in doing this kind of philosophy, one succeeds by using the best philosophical means available. One should not look for anything like a new recipe for doing so.

Second, although no conceptual engineer is guaranteed success, it still makes sense to engage in the systematic philosophical endeavor of criticizing concepts and attempting to improve or replace them. After all, no normative theory can be guaranteed success. There is e.g. a huge gap between building a theory that, if applied, would make our societies

[4] These objections are raised by Patrick Greenough in his *Against Conceptual Engineering* (unpublished manuscript).

more just and actually succeeding to build societies that correspond to the theory. While the theorist at least has some control over the former, the latter is surely beyond her control. But this does not mean that to engage in such normative theory building is senseless (Cappelen, 2018, ch. 7).

A third issue is called the Old Problem problem (Greenough MS) or the change of topic problem (Cappelen 2018). This problem was originally presented by Strawson (1963, p. 505) in response to Carnap's program of explication:

> It seems prima facie evident that to offer formal explanations of key terms of scientific theories to one who seeks philosophical illumination of essential concepts of non-scientific discourse, is to do something utterly irrelevant—is a sheer misunderstanding, like offering a textbook on physiology to someone who says (with a sigh) that he wished he understood the workings of the human heart.

The Old Problem problem arises for conceptual engineering precisely because of its aim to revise or replace the concepts by means of which the old problem was raised. Conceptual engineering thus appears to be a strategy which leaves the original problem unaddressed and can thus hardly be said to offer a solution to that very question, so the objection runs.

There are different ways of answering the Old Problem problem. One approach is to argue that if the old problem was raised by empty or inconsistent concepts, then surely, we do not even want an answer to *that* problem. Rather than a solution to that old problem, what we need instead are tools that let us see that the very question was misguided. Insofar as there are other kinds of conceptual deficiencies than emptiness and inconsistency, an alternative (or additional) approach is to argue that although the concepts have been altered by means of conceptual engineering, they can still be used to answer the old problems. This can either be because they play the same function or have the same role in our theories as the old concepts had[5] or because the topic under discussion can be preserved across the old and the new concepts

[5] This solution is argued for by Railton; for a presentation and criticism of his view see Cappelen (2018, ch. 10).

since topics are arguably coarser-grained than their accompanying concepts (Cappelen 2018, ch.10).

I believe this short survey of conceptual engineering and the contemporary debate provides sufficient background for the next two parts of the paper in which I will present my reading of Kant as a conceptual engineer.

Why Kant is a conceptual engineer

Let me begin by explaining why I take Kant to be a conceptual engineer. An essential feature of conceptual engineering, indeed the only feature that everyone in the debate appears to accept, is that conceptual engineering is a normative enterprise. It tells us how we *ought* to think, which representational devices we *ought* to use, how we *ought* to use them etc. That philosophy should be conceived of as a normative endeavor is also an essential part of Kant's meta-philosophy, i.e. philosophy as *critique*. And it is not only non-philosophers who fall prey to empty speak and get trapped in ambiguities, and thus need to be guided by such an endeavor. Indeed, Kant believes that all philosophers before him played around with empty words, lost themselves in paradoxes and landed on dogmatism or skepticism because they failed to uncover the proper concepts of their field, i.e. metaphysics. For although Kant argues that these concepts are a priori, they are not thereby easily accessible. A special transcendental exercise and methodology are required to lay them out and prove their status as proper metaphysical concepts. In other words, before metaphysics can proceed as a proper science, a critique which is a propaedeutic to that science must be established.

Kant is thus eager to distinguish his concern in the *Critique of Pure Reason* from that of Locke, whom he takes to ask merely the *quid facti*, i.e. how people in fact come to have and employ certain ideas.[6] Kant's own concern is instead the *quid juris,* i.e. the question concerned with what *right* certain representations are held to be, and can be employed as, genuine metaphysical concepts which give insight into the necessary properties of things (Kant 1998, A84/B116-A87/B199). If this right can be shown, the representation is not merely a concep-

[6] Whether this is a reasonable interpretation of Locke or not is beyond my present concern.

tion in someone's mind, which can be described by getting to know how people in fact think; rather, the concept that has been uncovered by Kantian means is an objectively valid representation. It reveals to one who grasps it what the properties of the object really are (albeit only *qua* appearance). Hence, it is this concept that people, including philosophers, *ought* to employ in their thinking. Insofar as we talk about concepts in this sense, the term "concept" is therefore a success term. But Kant does not only use "concept" in this sense. He also talks about concepts in a less demanding sense, cognitively speaking. In this latter sense, a concept is rather a mere thought or conception that may in different ways be wanting or defective. One example of the use of "concept" in this sense is when Kant talks about the concepts of pure reason such as <soul> and <world>, which according to the Transcendental Dialectic are illegitimate and rest on illusions and fallacious reasoning. Indeed, the *quid juris* itself arises precisely because of the suspicion that many of our philosophical concepts fall short of being concepts in the cognitively demanding sense. One kind of defect was suggested by Hume, namely that certain ideas, e.g. those of <cause> or <substance>, appear to be empty as no impressions correspond to them. This is what Kant (1900, Ak 4:260) has in mind when in the *Prolegomena* he famously states:

> I freely admit that it was the remembrance of David Hume which, many years ago, first interrupted my dogmatic slumber and gave my investigations in the field of speculative philosophy a completely different direction.

In the C*ritique of Pure Reason* Kant aims to ameliorate these ideas by means of his transcendental method. According to this method the concept proper is that which can be shown to function as a necessary condition for the possibility of experience. In doing so Kant does therefore not attempt to analyze what philosophers – or ordinary people for that matter – in fact mean by, e.g., "cause". Instead, he attempts to uncover what the metaphysical concepts must mean if they are to have the kind of objective reality that a successful metaphysical theory requires. The "must" here should be understood constitutively. The concepts he discovers are necessary conditions for the possibility of cognition. Nevertheless, his enter-

prise is normative as it guides and corrects how people, including philosophers, are prone to think as they fail to recognize the content and application conditions of the pure concepts of understanding.

The possibility of empty concepts is, however, not the only reason why philosophy must be conceived of as primarily a normative enterprise or a critique. In the Third Part of the *Prolegomena,* Kant (1900, Ak 4:338) says about cosmological ideas:

> This product of pure reason in its transcendent use is its most remarkable phenomenon, and it works the most strongly of all to awaken philosophy from its dogmatic slumber, and to prompt it toward the difficult business of the critique of reason itself.

From the point of view of conceptual engineering what interests me in this quote is how Kant's critique is an endeavor explicitly grounded in the recognition that we tend to form *inconsistent* concepts because, as we know, the cosmological ideas are concepts of reason that lead to antinomies, i.e. the assertion of both P and non-P. In other words, just like some of the contemporary conceptual engineers (Sharp and Eklund in particular), Kant holds that critiquing, ameliorating and sometimes replacing inconsistent concepts are among philosophy's most fundamental tasks.

Finally, for Kant there is also a third engineering task, namely that of detecting and subsequently splitting concepts that the philosophical tradition according to Kant has collapsed and hid under ambiguous terms such as "thing"; "inner"; "outer"; "substance"; "cause" and "logic". Uncritical talk about things should, e.g., be replaced by two distinct concepts, viz. that of <appearance> and <thing in itself>; the spatial inner and outer must be split from the non-spatial concepts of <inner> and <outer>; <substance> must be split from <subject> and <ground> from <cause>; and <logic> should be split into <general logic> and <transcendental logic>. Finally, on Kant's view some logical terms are also defective. As we shall soon see in some more detail, in opposition to the logic of his time, he argues that the concept of <negation> must be split into <not> and <non>.

So, to summarize, I take Kant to be a conceptual engineer because he holds that philosophy is a critical enterprise, it should not describe the concepts (in the cognitively weak sense) we happen to have, but rather replace or improve them. Improvement can take place by providing them with objectively valid content, by detecting and ameliorating inconsistencies or by splitting.

As I noted above, in the contemporary debate there are three issues of controversy in particular relating to: 1) the object of engineering, 2) the possibility of success and 3) the Old problem problem. Let me now briefly suggest how Kant's position as a conceptual engineer would fall, were he to be placed within this debate.

I think there can be little doubt what Kant's answer to the first issue would be. That which a philosopher critiques, improves or replaces are concepts, albeit understood in the cognitively weak sense when "concept" is not used as a success term. In this weak sense, even metaphysical concepts can be improved or replaced. On Kant's view, even though there are pure metaphysical concepts, i.e. the so-called pure *a priori* concepts of understanding, which he argues are constitutive of all cognition, the human mind is prone to misconstrue them, as can be witnessed by the failure of all previous attempts at metaphysics. Kant's transcendental method thus corrects our philosophical grasp or concept of the metaphysical concepts proper. By engineering concepts, one also changes the meaning of the corresponding word. But the word is not for Kant the primary target of engineering. Engineering of the concept also changes the world, partly because the concept's intension fixes the extension and partly because for Kant all conceptual amelioration aims in the end at moral, social and political improvement. Nevertheless, the primary object of conceptual engineering for Kant, what is worked at, is a concept.

Now what are such entities for Kant? Like some contemporary engineers (Sharp and Eklund), Kant conceives of concepts as constituted by principles which govern their application. More specifically the principles are judgments that can serve as the major premise in a certain kind of categorical syllogism. The judgments give you the *condition* under which the concept in question can be applied. Take for in-

stance the concept of substance. What we can call the *substance* principle is:

> What cannot be thought otherwise than as subject does not exist otherwise than as subject, and is therefore substance (Kant 1998, B411).

This judgment tells you that to be thinkable only as a subject (and not also as a predicate) is constitutive of a substance. Hence, <substance> should only be applied to objects that meet this condition.

Or take the *circle* principle:

> A line every point of which is the same distance from a single one (the center-point) [is a circle] (Kant 1998, A73/B759).

This judgement states the *condition* for something to count as a circle, namely that it must be a line whose every point is the same distance from the center point. This property is thus constitutive of circles and <circle> should only be applied to objects that meet the condition.

The minor premises in the corresponding syllogisms are judgments that make the claim that some object(s) meet(s) the condition of the major. Take for instance the following judgment: Matter cannot be thought otherwise than as subject. From these two premises the conclusion follows that matter is substance. Here is the categorical syllogism:

> What cannot be thought otherwise than as subject does not exist otherwise than as subject, and is therefore substance
>
> Matter cannot be thought otherwise than as subject
>
> ---
>
> Matter is substance

On the Kantian understanding of concepts, then, a concept can be wanting or defective in at least three ways: i) it can be empty in the sense that the relevant rule states a condition that no object in fact meets or even can meet, metaphysically or logically speaking; ii) it can be inconsistent because the relevant rule states a condition from which it follows that one and the same object both falls and does not fall under its con-

dition;[7] or iii) the condition of the relevant rule can contain ambiguous terms which conceal the need for splitting.[8]

Before I move on to a more detailed analysis of how Kant uses conceptual engineering to critique the dialectical inferences of reason, let me just say a few words about the two other issues in the contemporary debate. Recall first the challenge to the use of "engineering" as the label for this philosophical enterprise. The objection was partly that, as opposed to ordinary engineering, *conceptual* engineering has no genuine method, and partly that, because the meaning of language is fixed by external factors beyond our control, conceptual engineering can never succeed. As for the first of these worries, Kant thinks that there is at least one special philosophical method available for the improvement of metaphysical concepts, namely the method of showing that a certain concept is a necessary condition for the possibility of experience.

As for the worry concerning the lack of control, to address it properly one would have to delve into the complicated issue of comparing Kant's position to that of semantic externalism. For now, I will rest content with the answer that at least as far as the CPR goes, I think Kant would say that whether people will as a matter of fact grasp and apply the correct concept of, say, <infinity> after his engineering is beyond his concern. There will always be a gap between what a normative theory requires and what people in fact think and do even after they have achieved a good grasp of the theory. But the CPR offers at least some reason to hope that his engineering will not be in vain, for it is an important part of that theory that certain illusions arise naturally in the human mind, illusions which when undetected stand in the way of proper metaphysical insights. If Kant manages to uncover such illusions for his readers, one kind of obstacle in the way of the success of his endeavors is at least dealt with.

---

[7] This is the case e.g. with the concept of <infinity> in the first antinomy (Kant 1998, A427/B455).

[8] According to Kant at bottom most of the metaphysically defective concepts are characterizable by this kind of mistake, which he calls *Sophisma figurae dictionis* (Kant 1983, B411, B528). Sometimes these concepts are also contradictory in the sense explained above. Hence, ii) and iii) are not mutually exclusive conceptual defects.

The final issue that followers of conceptual engineering need to deal with is the Old Problem problem or the change-of-topic problem. This problem is also highly relevant for Kant's version of conceptual engineering. Indeed, it is for instance not *prima facie* unreasonable to object that Kant's purported solution to the question of how metaphysics is possible changes the very topic of metaphysics, as he replaces the concept of <thing> with that of <appearance>.

My brief answer is that Kant believes that if the revised concept plays at least partly the *same function* in the theory as the old concept did, we are entitled to argue that it can indeed address the old problem. So, for instance, although Kant's revised and improved concept of <cause> no longer covers mere logical ground or mere logical reason, as it did for the Leibnizians, Kant's concept of <cause> still plays the function of giving a metaphysical underpinning of the law-like behavior of objects and events in the world, and this was at least in part what the old concept of <cause> and its concomitant principle of sufficient reason were assigned to do. In some cases, however, where we deal with so-called dialectical concepts, Kant would be happy to dismiss the old problems as he argues that they are grounded in metaphysical illusions. In such cases we should opt for replacement rather than revision of the original concepts, for what we need in these cases is precisely not to answer the old problem, but to see that the question itself, raised by the old defective concepts, is misguided.

Kant's *Critique of Pure Reason* – a piece of conceptual engineering

My aim in this part of the paper is to argue that Kant's critique of the arguments of speculative metaphysics, as presented in the Transcendental Dialectic, can be interpreted as resting on a piece of conceptual engineering.[9] However, to

[9] This reading is radically different from the standard readings according to which Kant's arguments in the Transcendental Dialectic are mainly therapeutic for we already know from the Transcendental Analytic that we cannot achieve knowledge of such transcendent objects as the soul, the world as a totality and God. It falls beyond the aim of this paper to argue against these readings.

show this for all the arguments would be impossible within the limits of this paper.[10] To make my case I will therefore focus on his critique of two of the arguments presented in the chapter on the paralogisms. In this part of the Dialectic, Kant's target is the inferences of the rational psychologist, which purportedly yield insight into the properties of the soul, conceived of as an object that transcends the bounds of space and time. As we shall see, Kant's reconstruction of the inferences shows that the proponents of these apparently valid inferences are in fact caught up in logically invalid reasoning. This is concealed to them due to an ambiguity in the term "not". The ambiguity obscures two different formal concepts, namely <not> and <non>, which need to be split and kept strictly apart as their truth conditions differ. My reading thus brings out that Kant's dismissal of the arguments of the rational psychologist rests upon his critique and engineering (in the sense of splitting) of the concept of negation into two distinct concepts. To grasp this splitting we first need to turn to Kant's logic.

Kant conceives of logic as a set of rules for the acts of thinking:

> Logic is the science that exhaustively presents and strictly proves nothing but the formal rules of all thinking (whether this thinking be empirical or priori, whatever origin or object it may have, and whatever contingent or natural obstacles it may meet with in our minds) (Kant 1998, Bviii-ix).

Thinking for Kant is judging: he calls the acts of thought "judgments" and argues that there are twelve different logical forms of judgment which cannot be further reduced (Kant 1998, A70/B95). For my purpose in this paper, only two of these forms are relevant, namely the two forms of negation called negative and infinite judgment.

Kant describes the logical form of negative judgment in the following two-fold way: i) in a negative judgment the negation affects the copula (Kant 1900, Ak 9:104, n.3; Ak 24:76) and ii) "[I]n the negative [judgment] the subject is posited outside the sphere of the [predicate]" (Kant 1900, Ak 9:104). In other words, a negative judgment commits one to a claim

[10] For a richer presentation of this interpretation see Serck-Hanssen (2019).

about what the subject *is not*, and thus removes the subject from the domain of the predicate in question. But it does not thereby posit it in another domain. To put the point in less Kantian jargon, the truth of a negative judgment has no existential import and is compatible with empty subject terms as it says no more than "it is not the case that S is P," which is trivially true if "S" is empty.

An infinite judgment, though it is also a kind of negation, is subject to a quite different logical form. To be entitled to use this form one must not only remove S from the sphere of P, S must be posited *in* a sphere, namely the infinite one outside the sphere of the predicate that the judgment negates. Unlike negative judgment, infinite judgment is thus an act that preserves the claim that S *is* something. It says in effect that S *is* non-P. Accordingly, both in the *Jäsche Logic* and in the *Dohna-Wundlacken Logic*, Kant says that in an infinite judgment it is not the copula but the predicate that is affected by the negation (Kant 1900, Ak 9:104, n.3; Ak 24:764).

The same point is made in the Vienna logic (Kant 1900, Ak 16:930):

> *Anima non est mortalis* is a negative proposition. On the other hand, *Anima est non mortalis* is an infinite proposition…Affirmation and negation are qualities in judgment, accordingly. A negative judgment is not just any judgment that is negative, but a negative judgment where the negation affects the *copula.* A judgment is an affirmative judgment, accordingly, where it does not affect the *copula* but rather the predicate, as occurs in an infinite judgment, and where the *copula* is without any negation[;] consequently, all infinite judgments are affirmative, because the negation affects only the predicate. But although every infinite judgment has the nature of the affirmative, nonetheless, there is always a negation there, not of the judgment, i.e., of the relation of the concepts, but of the predicate. The relation is the same, to be sure, as in an affirmative judgment, but the negation is still always there, and consequently infinite judgments are distinct from the affirmative judgment. In logic, this matter seems to be a subtlety. But in metaphysics it will be a matter of importance not to have passed over it here.

Since an infinite judgment negates the predicate whereby it posits the subject in a domain, an infinite judgment is necessarily false if the subject term is empty.

Now, the idea that there are two kinds of negation is not new with Kant. We find infinite negation already in Aristotle, and in medieval philosophy the logic of infinite negation became a topic of great interest as they discussed the proper interpretation of the O form (Some S is not P) in the traditional square of opposition. A number of logical principles for contraposition and obversion were suggested to allow one to move between so-called simple and infinite negation (Parsons, 2017). The great medieval logician Buridan showed, however that many of these principles failed in the case of empty subject terms. One of his funny examples was the move from the (true) negative judgment "A chimera is not a man" to the (false) infinite judgment "A chimera is a non-man".

Buridan's insights ruled for several hundred years and logicians realized that unless one added a premise that precluded empty subject terms one could e.g. not move from the truth of 'Some S is not non-P' to 'Some S is P'. But somehow, the subtle discussions and insights of medieval logic were lost. From the middle of the sixteenth century onwards, spurious principles of contraposition and obversion were back on the scene, without any sign of worry about empty terms. As Parsons notes (2020), the highly influential Port Royal Logic's "discussion of the O form is so vague that nobody could pin down its exact truth-conditions, and there is certainly no awareness indicated of problems of existential import." One of the invalid principles that made its way back was that 'Some S is P' is equivalent to 'Some S is not non-P.' As we shall soon see, according to Kant's reconstruction this invalid principle is used by the proponent of speculative metaphysics when attempting to establish that the soul is simple.

Kant had to teach an excerpt of Meier's logic for almost 40 years. Like those of the rest of the logicians in that period, Maier's logic shows no awareness of the problem of moving from simple to infinite negation. Here are two representative quotes:

> If, in particular negative judgments one applies the negation to the predicate, then they become particular affirmative judg-

> ments [§294], and then they can be converted absolutely (Meier 2016, §351, p. 79).
>
> Thus, one can transform all negative judgments into affirmative ones, if one moves the negation from the concept of combination to the predicate. E.g. the soul is not mortal, the soul is immortal (Meier 2016, §294, p. 68).

As far as I know Kant never explicitly criticizes Meier for having proposed invalid rules. Nevertheless, his own lectures on logic show that Kant himself avoids making the same mistakes. As we have just seen, Kant takes the infinite form of judgment and the negative form to have different truth conditions and his example in the Vienna logic is precisely the same as Meier's. Pace Meier, one therefore cannot transform all negative judgments into affirmatives (i.e. what Kant calls infinite judgments) by moving the negation from the concept of combination (i.e. the copula) to the predicate. Hence, the truth of "the soul is not mortal" does not logically imply the truth of "the soul is non-mortal (i.e. immortal)". The latter only follows if one assumes that "soul" is not an empty term. And this assumption is – to borrow Kant's words from the Vienna logic – no subtlety but a matter of importance in metaphysics, for here the existence of such objects as the soul which allegedly transcend the bounds of space and time is just what is under dispute. Hence, none of the parts in this metaphysical debate can allow themselves to simply presuppose that their subject term "soul" is not empty.

Kant's distinction between infinite and negative judgment and the corresponding concepts <non> and <not> is a piece of conceptual engineering. In arguing for the distinction, he lays down rules that split the concept of negation into two distinct concepts with different truth conditions. As we shall now see, this splitting allows him to reveal that apparently flawless arguments of the speculative metaphysician are in fact logically fallacious.

As my first example of how Kant makes use of his engineering of the concept of <negation> in his critique of speculative metaphysics, I have chosen the first paralogism as presented in the B-edition of the CPR. The inference runs like this:

> What cannot be thought otherwise than as subject does not exist otherwise than as subject, and is therefore substance.
>
> Now a thinking being, considered merely as such, cannot be thought otherwise than as subject.
>
> ---
>
> Therefore, it also exists only as such a thing, i.e., as substance. (Kant 1998, B410–411)

According to Kant this inference is formally invalid. It allegedly commits the mistake of the kind *sophisma figurae dictionis*, in which one operates with ambiguous middle terms (B411). Thus, despite the appearance of being a valid three-term categorical syllogism of the form:

All S are P

a is S

---

a is P

Kant seems to hold that the form of the paralogism is the invalid four-term kind:

All S are P

a is R

---

a is P

In contemporary logic, a fallacy of equivocation would normally not be thought of as a *formal* fallacy since it depends on the meaning of terms, and several commentators have argued that Kant is simply wrong: formally speaking, the inference is perfectly fine. However, as we shall now see, the peculiar character of this inference is that the ambiguity of the middle term can be construed as pertaining to a logical term, namely the term "not".

In the categorical syllogism rendered above, the middle term is "what cannot be thought otherwise than as subject". This term appears as the subject term in the major premise

and as the predicate in the minor. It thereby functions as the medium or ground that allows one to draw the conclusion in which the minor term ("a thinking being considered only as such") and the major term ("substance") become connected.

At first glance, there appears to be no ambiguity here. Notice, however, that this middle term involves the use of negation. In addition to the explicit negation of possibility (of thinking something), "what *cannot* be thought", there is also another negation, namely the one pertaining to the way the object (referred to by "what") is thought, for "otherwise than as subject" is just another way of saying "not as subject". Now, in this context, this is equivalent to "as predicate". By combining the two negations in the original middle term, we get middle term*: "what must be thought not as predicate". This middle term is ambiguous, for as we saw above, Kant would hold that the word "not" is ambiguous. It can refer to two different concepts of <negation>, viz. <not> and <non>, which are governed by the rules of negative and infinite judgment, respectively.

We are now in a position to see how Kant's splitting of the two concepts of negation gives him a powerful tool against the proponent of the apparently flawless inference about the soul. Notice first that, given the split between <non> and <not>, the middle term can be interpreted in two ways. It can either be read as "what must be thought as a non-predicate" or it can be read as "what must be thought as (it is) not a predicate". Now, in the major premise the middle term should clearly be read as expressing the infinite kind of negation. As we have seen, if a judgment of this form is to be true, which the major premise is, according to Kant, it presupposes non-empty subject terms. And indeed, Kant makes this condition explicit by adding to the major premise "then it *exists* also as . . .". In the minor premise, however, although the rational psychologist might believe that he has the right to use the middle term in the same sense as in the major, this is not correct, according to Kant. The thinking being about which it is true to say that it must be thought as not a predicate is merely the logical I of the principle of apperception (B407). But from this I, no existence claim can be inferred. Thus, on the basis of this I alone, which according to Kant is the "sole text" of rational psychology (A343/B401), the middle term in

the minor premise must be read as instantiating mere negation: "what must be thought as (it is) not a predicate". This reasoning yields the following version of the inference:

> What must be thought as a non-predicate does not exist otherwise than as subject, and is therefore substance.
>
> Now a thinking being, considered merely as such, is what must be thought as (it is) not a predicate.
>
> ---
>
> Therefore, it also exists only as such a thing, i.e., as substance.

Since the form of the minor premise is a negative judgment, it can be true even if the subject term is empty. This shows that this inference is logically invalid and that the desired conclusion does not follow. Moreover, pace Meier the minor premise cannot be transformed into an infinite judgment either, for there is no valid logical route from *S is not P* to *S is non-P*. Kant's splitting of the concept of negation thus succeeds in disclosing that one of the most convincing arguments of speculative metaphysics fails.

My second example is also taken from the paralogisms, and this time I turn to the so-called Achilles argument of the second paralogism in the A-edition. The argument follows in the prose just after the presentation of the syllogism which purports to prove that the soul is simple:

> That thing whose action can never be regarded as the concurrence of many acting things, is simple.
>
> Now the soul, or the thinking I, is such a thing.
>
> ---
>
> Thus etc. [the soul or I is simple] (A351)

Here is how Kant presents the *Achilles argument:*

> This is the Achilles of all the dialectical inferences of the pure doctrine of the soul, nothing like a mere sophistical play that a dogmatist devised in order to give his assertions a fleeting plausibility, but an inference that seems to withstand even the sharpest testing and the greatest scruples of inquiry. Here it is.

> Every composite substance is an aggregate of many, and the action of a composite, or of that which inheres in it as such a composite, is an aggregate of many actions or accidents, which is distributed among the multitude of substances. . .Yet with thoughts, as accidents belonging inwardly to a thinking being, it is otherwise. For suppose that the composite were thinking; then every part of it would be a part of the thought, but the parts would first contain the whole thought only when taken together. Now this would be contradictory. …Thus it is possible only in one substance, which is not an aggregate of many, and hence it is absolutely simple (A352).

As Van Cleve has noticed (1999, p. 176), this suggests that the syllogistic inference Kant actually criticizes is not the syllogism as first presented, but rather the Achilles version in which the original major premise is replaced by Major* by means of contraposition:

> Major*: If something is non-simple (i.e. composite), its action can be regarded as the concurrence of many acting things.

The next part of the argument says that it is contradictory to assume that the action of the soul, viz. its thinking, is an action of a composite. This line of reasoning gives us the minor premise of the inference:

> Minor: The action of the soul cannot be regarded as the concurrence of many acting things.

From this we can draw the conclusion by modus tollens:

> Conclusion: The soul is not non-simple (i.e. composite).

So far, so good. The problem is, however, that this conclusion is not what the rational psychologist wants. He wants to show something more, to wit that the soul is simple. Although "the soul is not non-simple" is true (if the premises are true, and Kant says they are), and it is indeed vacuously true in the case that "soul" is empty, this truth is of little interest to the metaphysics of the soul.

To prove his case the speculative metaphysician might attempt to move from the negative conclusion obtained through this chain of reasoning: "it is not the case that the soul is non-simple", to an affirmative judgment: "the soul is

simple". However, as we already know, by splitting the concept of negation, Kant was brought to the insight that such moves are formally illicit. The vulnerable heel of Achilles, the apparently flawless master argument of speculative metaphysics, is thus disclosed by a piece of conceptual engineering.

Notice finally that Kant's critique of the Achilles argument does not imply that "soul" is an empty term; the point is rather that the burden of proof is now on the rational psychologist. It is he who will have to come up with a theory showing that "soul" does indeed have reference to the kind of object that the theory requires, namely an object beyond the bounds of space and time. Moreover, until that theory is at hand, the rational psychologist cannot respond like Strawson did to Carnap: that Kant's solution fails to answer *his* questions about the soul. If Kant's transcendental idealism is correct, the rational psychologist's questions are indeed misguided, for the purported objects of speculative metaphysics are no more than objects of thought (*Gedankendinge*) or nothing (*ens rationis*) in the sense of empty concepts without objects (A490/B517, A543/B571, A566/B594, A292/B348, A681/B709). And in that case the rational psychologist's concept of soul should be replaced rather than revised.[11]

Conclusion

I have tried to make the case for reading Kant as a conceptual engineer. I have argued that his normative understanding of philosophy and his preoccupation with critiquing and improving possibly empty, contradictory and ambiguous concepts make this label appropriate. I have also suggested how Kant would answer some of the challenges that conceptual engineering faces, such as the question of what the object of engineering is, what concepts are and the Old Problem prob-

[11] Notice, however, that there are other concepts of soul that Kant embraces. He operates both with an empirical concept of soul and also with a practical concept of soul, which plays a crucial role in his moral philosophy. The concept of soul that I argue must be replaced is that which is generated by the search for the unconditioned, a concept which according to Kant is inextricably tied up with illusions and the faulty position of transcendental realism.

lem. In the third part I have attempted to show both how Kant splits the concept of <negation> into <not> and <non> and how this splitting can be used to show that inferences of speculative metaphysics are fallacious. Kant's revelation that the purported science of rational psychology is grounded in a failure to split and keep apart two different concepts of <negation> is on my view the sign of a conceptual engineer par excellence.

University of Oslo

References

Burgess, A. and Plunkett, D. (2013a), "Conceptual Ethics I", *Philosophy Compass*, 8 (12), pp. 1091-1101.

Burgess, A. and Plunkett, D. (2013b), "Conceptual Ethics II", *Philosophy Compass*, 8 (12), pp. 1102-1110.

Cappelen, Herman (2018), *Fixing Language: An Essay on Conceptual Engineering*. Oxford: Oxford University Press.

Eklund, M. (2014), "Replacing Truth?" In: Burgess, A. and Sherman, B. (eds.) (2014). *Metasemantics: New Essays on the Foundations of Meaning*. Oxford: Oxford University Press.

Eklund, M. (2017), *Choosing Normative Concepts*. Oxford: Oxford University Press

Greenough, P. (MS), *Against Conceptual Engineering*.

Haslanger, S. (2012). *Resisting Reality: Social Construction and Social Critique*. Oxford: Oxford University Press.

Kant, I. (1998), *Critique of Pure Reason*. Translated and edited by Paul Guyer and Allen Wood. Cambridge: Cambridge University Press. (Cited, in standard A/B edition pagination with 'A' and/or 'B' and page numbers following.)

Kant, I. (1900–), *Gesammelte Schriften*. Kǒnigliche Preussische (later Deutsche) Akademie der Wissenschaften (Ed.). Berlin: Georg Reimer, later Walter de Gruyter. (Cited as Ak., followed by volume and page number. All translations are from the Cambridge edition.)

Kjosavik, F. and Serck-Hanssen, C. (eds.) (2019), *Metametaphysics and the Sciences: Historical and Philosophical Perspectives*. London: Routledge.

Leslie, S.-J. (2017). "The Original Sin of Cognition: Fear, Prejudice, and Generalization". *Journal of Philosophy*, 114 (8), pp. 393-421

Meier, G. F. (2016). *Excerpt from the Doctrine of Reason*. Translated by A. Bunch. Bloomsbury, London.

Parsons, T. (2020), "The Traditional Square of Opposition" *(Stanford Encyclopedia of Philosophy/Summer 2017 Edition)*, available online at: https://plato.stanford.edu/archives/sum2017/entries/square/ [Accessed 2 Jan. 2020].

Plunkett, D. (2015), "Which Concepts Should We Use?: Metalinguistic Negotiations and The Methodology of Philosophy", *Inquiry: An Interdisciplinary Journal of Philosophy*, 58 (7-8), pp. 828-874.

Plunkett, D. and Cappelen, H. (2020), "A Guided Tour of Conceptual Engineering and Conceptual Ethics" in Cappelen, H. and Plunkett, D. and Burgess, A. (eds.), *Conceptual Engineering and Conceptual Ethics*, Oxford University Press, Oxford.

Serck-Hanssen, C. (2017), "Fighting Achilles", in Motta, G. and Thiel, U. (eds.) *Immanuel Kant: Die Einheit des Bewusstseins*, De Gruyter, Berlin, pp. 130–147.

Serck-Hanssen, C. (2019), "From Nothing to Something – Why Metaphysics Cannot be Reduced to Logic", in Kjosavik, F. and Serck-Hanssen, C. (eds.), *Metametaphysics and the Sciences: Historical and Philosophical Perspectives*, Routledge, London, pp. 134-152.

Scharp, K. (2013), *Replacing Truth.* Oxford University Press, Oxford.

Sterken, R. (2020), "Linguistic Interventions and Transformative Communicative Disruption". In: Cappelen, H. and Plunkett, D. and Burgess, A. (eds.), *Conceptual Engineering and Conceptual Ethics*. Oxford University Press, Oxford.

Strawson, P.F. (1963), "Carnap's Views on Constructed Systems vs. Natural Languages in Analytic Philosophy", in Schilpp, P.A. (ed.), *The Philosophy of Rudolf Carnap*, Vol XI, Open Court, La Salle, IL.

Van Cleve, J. (1999), *Problems from Kant*, Oxford University Press, New York.

Logical and Epistemic Normativity

On the Normativity of Logic

MIRCEA DUMITRU

Modal and normative concepts are plenitudinous.[1] Here there are some of the most often used concepts in philosophical explanations and investigations: the necessary truths of logic, mathematics, and metaphysics, respectively; the necessary connections among events or states of affairs in the natural world; the necessary or unconditional principles of ethics; and many other forms of necessary truth or connection, forms of oughts, of should a.s.o.

It is quite reasonable, then, that this abundance raises the legitimate question: how much *real* diversity are we really having here? Are all those modal and normative concepts independent from each other, or may they be reducible to just one kind or to very few kinds of primitive modal concepts? And if the latter, are there some irreducible ways in which a truth might be necessary or a connection might hold by necessity?

The leading concept which is constitutive for my position here regarding the normativity of logic — or at least the way in which I make sense of this topic — is an anti-reductionist stance according to which there are three main forms of necessity, viz. the metaphysical, the natural, and the normative; and each of them is irreducible to the others or to any other form of necessity.[2]

Now, logic itself is replete with modal/intensional/normative concepts; one regularly frames the logical/metalogical

[1] I would like to express my gratitude to two anonymous referees for making remarks which have contributed to the improvement and clarification of some parts of my paper. It goes without saying that I am the only one responsible for the contents and arguments of the paper.

[2] Cf. K. Fine 2006, p. 235.

notions in terms of necessary truths, connections, etc., or in terms of what should be the case or what shouldn't be the case, of what ought to be the case a.s.o.

In what sense, or senses, can logic be said to be normative? For sure, in a sense, this seems to be a trivial issue. For logic *appears* to be the paragon of a normative science. The feeling that logic typifies in a very strong sense the very concept of a normative science which is to be sharply opposed to descriptive and empirical inquiry is so robust that it motivates the further question of whether logic might enjoy an *exceptional status* amongst the other sciences.

To make our way into the topic a guided tour through the complex landscape of the literature is worthwhile. To begin with, it pays off to discuss the thesis of the normativity of logic, which is the thesis that logic exerts a normative constraint on reasoning, within the current debate between *logical exceptionalism* and *logical anti-exceptionalism*, respectively. In recent years there has been a growing literature concerning the epistemological, metaphysical and methodological status of logic, with a particular focus on the contrast between exceptionalists and anti-exceptionalists.[3]

The exceptionalist view

The *standard view* about logic — at least within the analytic tradition[4] — is that logic is an *a priori* science. Thus, empirical evidence doesn't have — and, even more strongly, *cannot* have — any significant connection with the justification of logic.[5] Many logical apriorists contend that logical principles and inference rules are justifiable *a priori* and, consequently, they are empirically indefeasible.

---

[3] For making sense of the intricacies of the conceptual landscape into which the controversy between exceptionalism and anti-exceptionalism is rooted I benefited a lot from the research of Filippo Ferrari, and Filippo Ferrari and Sebastiano Moruzzi and Florian Steinberger. My own paper mainly consists in a reaction, of the kind of lecture notes and comments upon their illuminating work.

[4] Carnap 1937, Dummett 1991, Frege 1893, Rumfitt 2015, Tarski 1936, 1956, Wright 1986.

[5] Field 1996, 1998, 2005.

Logical apriorists picture logic as merely *formal*.[6] Consequently, logic studies what follows from what in an abstract, general and topic-neutral way, which makes logic independent of factual issues pertaining to the nature of the empirical world. This independence of logic from the world also makes the former immune from revision on empirical grounds.

In addition to all this, in contradistinction to the empirical sciences, logic is also *a priori* because it is pervasive, fully general, and fundamental for thinking. The principles and inference rules of logic are constitutive for the general structures of any kind of human reasoning and argumentation. They are the infrastructure of scientific reasoning and of philosophical argumentation as well.

So, for example, in doing *empirical science* one uses logical rules when making inferences from empirical observations. In *mathematics* the use of logical rules is fundamental for proving mathematical theorems. *Logic* itself makes use of logical rules when it comes to evaluating logical relations between propositions. More to this point, logic itself is required when logic is the object of research; and likewise, when aiming at figuring out the kind of methodology for revising a logical system we make use of logic again.[7] So, in principle, there cannot exist a logic-free appraisal of logic. And in more recent times, *metaphysics*, especially in the analytic tradition, makes heavy use of logic to fathom models of the hidden structure of reality and its modal metaphysical grounds.[8]

So, for some prima facie good reasons, logic has been conceived of as belonging in the family of *normative* disciplines, which makes it kind of analytical that logic is in the business of producing *norms* for thinking and reasoning. And the fact that the methodology of logic should be considered as being in *strong contrast* with that of the empirical sciences is a natural consequence of logic being a normative, i.e., non-descriptive and non-factual, discipline. Some philosophers go even further and contend that logical norms are grounding the thinking and the reasoning to the extent that whoever

---

[6] Tarski 1936.

[7] Dummett 1991, Woods 2018, 2019.

[8] Fine 2012, Correia & Schnieder 2012, Timothy Williamson 2013.

does not follow the laws of logic is not a genuine thinker and reasoner any more.

Kant's stance exemplifies very well this normative concept of logic, for what he calls "pure general" logic "contains the absolutely necessary rules of thought without which there can be no employment whatsoever of the understanding".[9] Gottlob Frege contends that the logical laws "are the most general laws, *prescribing* (italics mine) how to think wherever there is thinking at all".[10] Now, there are philosophers who interpret Frege's account of the *normativity of logic* as leading to *the constitutivity thesis* of logic for thought;[11] however, there are some other philosophers who argue that according to Frege *illogical thought is possible*, but it would be *incorrect* and *irrational*. Regardless the different interpretations, one thing is sure: Frege teaches us that logic is aiming at the laws of truth: "to discover truth is the task of all the sciences; it falls to logic to discern the laws of truth".[12] Moreover, since reasoning, judging and inferring seek the truth, the laws of logic will govern all those thought operations, thereby turning the laws of logic, which are objective and indefeasible, into laws of thought. Wittgenstein makes similar points when he contends that logic being constitutive for thought establishes its *a priori* character: "What makes logic *a priori* is the impossibility of illogical thought".[13]

The epistemology of logic, and in particular the issue concerning the normative character of truth and meaning, have attracted more attention lately. The leading concept around which the *standard view* of logic has been developed was the concept that logic is by its very nature a normative discipline which falls into the same family with ethics. What ethics is to the normative constraints on proper human conduct, logic is to the normative constraints of truth, meaning and reasoning. This is the core of the view that conveys to logic an *exceptional status* among the sciences, belonging into the family of normative disciplines rather than into that of empirical inquiry.

---

[9] Kant 1781, A52/B76.
[10] Frege 1893, p. xv.
[11] Steinberger 2017.
[12] Frege 1956, p. 289.
[13] Wittgenstein 1921, §5.4731.

The whole dialectics which is going on in this literature has motivated an interesting twist as a result of subjecting the whole idea that logic is a normative science to a thorough scrutiny and criticism connected with a line of thought articulated by many authors.[14] The impact of those discussions resulted in a renewal of the interest in the topic of the normativity of logic stemming especially from philosophers of logic preoccupied with *epistemic normativity*.

The anti-exceptionalist views

Opposed to the standard view, according to which logic is an *a priori* and empirically unrevisable discipline which provides the most fundamental and indispensable rules of thought that normatively constrain our reasoning and judging, there are various accounts of logic that consider the methodology of logic to be *broadly abductive*.

Those accounts thus allow for some extra-logical considerations to impact on the legitimacy of fundamental logical principles. And this goes by either (i) considering logic to be continuous with the empirical sciences, and thus open to revision on the basis of empirical observations,[15] or (ii) otherwise answerable to general abductive considerations primed by theoretical and explanatory virtues such as simplicity, unification, explanatory power, etc.[16]

Anti-exceptionalism about the methodology of logic consists precisely in this *abductive* and, arguably, *non-aprioristic* view of logic (although some authors are considering the possibility of *a priori* abduction[17]). The main tenet of this stance is that logic does not have a special status among the sciences: "its methods are continuous with scientific method. Logic isn't *a priori*, nor are its truths analytic truths. Logical theories are revisable, and if they are revised, they are revised on the same grounds as scientific theories".[18] There is this well-known case which exemplifies strategy (i) above, which is

---

[14] Field 2009, Pedersen 2019, Russell 2017, Steinberger 2017.

[15] Quine 1951, Putnam 1979.

[16] Bueno 2010, Bueno & Colyvan 2004, Priest 2006a, 2006b, 2014, 2016, Williamson 2013, 2017a, 2017b, Hjortland 2017, Russell 2015.

[17] Biggs & Wilson 2017.

[18] Hjortland 2017, p. 632.

due to Quine and Putnam. They argued that some empirical findings and interpretations of quantum physics may very well motivate a revision of some fundamental principles of classical logic, such as for instance the distributive law of the conjunction over the disjunction, namely the law that licenses us to infer '(p & q) V (p & *r*)'from 'p & (q V *r*)'. As a result, according to Quine and Putnam, a revision — in fact, an abandonment — of classical logic in favor of quantum logic was thought to be required. This idea that logic is revisable on the basis of empirical consideration is directly in tension with the aprioristic conception of logic described above under the rubric of the standard view of the aprioristic methodology of logic.

According to anti-exceptionalists there are *different kinds* of evidence which can determine the legitimacy of a logical theory. In addition to empirical evidence, there are both *a priori* and *a posteriori* considerations which might ground the adequacy of a logical theory, and what we mean here are metaphysical, semantic, epistemic, linguistic or mathematical considerations.

I shall give you a set of short comments about how those considerations enter into the justifications of various brands of logics. *Intuitionistic logic* gets an account in terms of metaphysical and semantic considerations which reject the assumptions underlying metaphysical realism and support a constructivist concept of verifiability instead of classical truth. Thus, M. Dummett, who defends intuitionistic logic, construes the debate between realism and anti-realism as a controversy about whether or not the nature and the role of the truth-conditions are constitutive for the meanings of the sentences of a given discourse. An account of the meaning of a sentence through its truth-conditions would be satisfactory provided it also gave a plausible account of what the understanding of that sentence consists in. However, Dummett argues that the classic semantic account of the notion of truth and, moreover, the whole concept of understanding which realism seeks to ground run into serious troubles. This is how, according to Dummett's strategy, the old metaphysical debate concerning realism and truth should be conducted in the field of the philosophy of language; and correspondingly, an epistemic account of truth should be substituted for the

classic semantic account of truth.[19] *Relevance logic* seeks to capture in a more adequate way our common linguistic intuitions concerning the use of implication in everyday reasoning as asking for a relevant content shared by both the antecedent and the consequent of the implication. *Fuzzy logic* aims at building an adequate theory which will give us a thorough understanding of vagueness. And last, but not least, *paraconsistent logic* looks into the intricacies of inconsistent theories which are non-trivial and fruitful from an epistemic standpoint.

Then again, the way we determine how we make our choice of a logical theory does not differ essentially from how we make a theory choice in some other more empirical sciences. And this is the case because of a set of *general epistemic criteria* which operate in a similar manner in logic and in the other sciences. Here there are some of those important criteria for theory choice which operate across the whole spectrum of sciences: simplicity, explanatory and unificatory power, coherence, non-adhocness, systematicity, empirical adequacy and independent testability.

This epistemic belief of the anti-exceptionalists in the smooth operations of roughly the same criteria of theory choice across all the sciences, including logic, further motivates their belief that logic, just as any other scientific discipline, should make use of abductive reasoning when one makes a theory choice or a theory revision. So, when choosing a logical theory one is supposed to keep in mind what theory gives the best explanation for a given logical puzzle or problem; and also, one is supposed to balance the various abductive criteria.

Opposing the received view of logical apriorism, anti-exceptionalism including this distinctive abductive methodology has become quite a fashionable trend in contemporary philosophy of logic, or more specifically epistemology of logic.

Anti-exceptionalists have intensely debated some issues of genuine interest regarding the ways abductive criteria are supposed to be used in the process of logical theory choice. I shall mention some of them only: how are we going to identi-

---

[19] Cf. M. Dummett, 1978, 1991.

fy those abductive criteria which are relevant for logical theory choice; in what way or ways can we give relative priority to each such criterion, since different ways of weighing the criteria will result in different rankings of the logical theories we assess;[20] what is the role which the background logic should play in assessing the methodological and epistemic merits of the logical theories that are answerable to the logical theory choice.

Some anti-exceptionalists go all the way through with their stance about logic not being exceptional amongst the other sciences. Accordingly, they insist that the background logic which is the springboard for the logical theory choice and assessment needs not be an indefeasible *a priori* logic. Even the background logic may be chosen through the application of certain *abductive criteria*. Thus, O. Bueno[21] recommends a very liberal stance towards adopting a background logic in a kind of non-a priori way: *accepting* that logic is just enough as a prerequisite for starting the abductive machinery of assessing and making one's choice of a logical theory. Moreover, the accepted background logic might be subject to a revision procedure in another context. Likewise, Mark Sainsbury, embraces a holistic perspective on the choice or revision of a logic theory when he urges us to adopt in logic Neurath's boat metaphor for the empirical sciences: "The logical boat, as much as the empirical one, must be reconstructed at sea, plank by plank, with no privileged and with no pre-ordained right starting-point".[22]

Nevertheless, the background logic we may use in our process of logical theory choice may influence in various degrees our consequential decision as to how well a logical theory satisfies the abductive criteria and how well the chosen theory scores compared to other logical theories.[23] Here I also hint at important abductive benchmarks such as coherence, avoidance of adhocness or explanatory power, which are contextually defined through the concept of *logical inference of the background logical system*. And here we are in a sort of predic-

[20] Routley 1980, Priest 2016.
[21] Bueno 2010.
[22] Sainsbury 2002, p. 14.
[23] Cf. Woods 2018, 2019.

ament, since, as we already pointed out, there is no logic-free standpoint from which we may abductively assess the pros and cons of a logical theory. So, what we may very well end up with through the application of those abductive criteria is a situation in which a given logic L1 may result in different scores when assessed from the perspective of a background logic L2 than when again assessed from the perspective of a different background logic L3. Priest very cogently contends that "[a] worst-case scenario is one where we simply flip back and forth between two logics [...], each of which is better according to the other!"[24]

Consequently, the anti-exceptionalists face this challenge: how not to be trapped into "the problem of revision cycles"[25] when assessing logical theories within a given background metalogical theory. And this problem leads further to the more general well-known difficult situation of what background meta-theory one should adopt when defending a given logical system. A classic logician can stick consistently to the same classic logic framework both at the metalogic level and at the level of the logic which she argues for. However, a non-classic logician (who may be a devotee of a non-classic deviant logic whatsoever and of logical pluralism) is in an alarming position for she either stays classic in the metalogic background, which will defeat her arguments in favor of her preferred non-classic logic or else she will consistently embrace the non-classic logical theory at both levels, i.e. logic and metalogic, which will make her the target of the criticism of buying into a sort of circular justificatory procedure for her own preferred logic.[26]

One other hot issue in current philosophy of logic is the ongoing debate between *logical monism* (the view that there is only one correct logic) and *logical pluralism* (the view that there is more than one correct logic). It is worthwhile pointing out that the disagreements between the fans of either camp are feeding correspondent disagreements among anti-exceptionalists about whether abductive methodology in logic supports *logical monism* or *logical pluralism*.

---

[24] Priest 2016, p. 17.
[25] Cf. Woods 2018.
[26] Sereni & Sforza-Fogliani 2017.

Graham Priest and Timothy Williamson are two outstanding logicians-philosophers who side with the *anti-exceptionalist* conception in their corresponding philosophy of logic. They both share the view that the process of logical theory choice should be based on evidence from a truth theory qua an essential abductive criterion. Interestingly enough, though, even if they both argue from there to a conclusion supporting logical monism, each one backs monism in his own way which is completely divergent from the other's way. Thus, Priest contends that an anti-exceptionalist stance leads to some brand of paraconsistent logic.[27] And as opposed to this non-classic position, Williamson forcefully argues in favor of classical logic.[28]

This is the right place where the following caveat seems to be in order. My point is not polemic; rather, what I am interjecting here is in the spirit of a clarification remark. There is this terminology which Florian Steinberger uses and which has been taken up by Ferrari and Morruzi, viz. that of *exceptionalism vs. anti-exceptionalism*. In many ways this dichotomy seems to coincide with the *monism vs. pluralism* dichotomy.

Nevertheless, the two dichotomies do not coincide; for, as I have already pointed out, one can be an exceptionalist and a pluralist, which is a combination absolutely in order, but one can also be an anti-exceptionalist and a classicist, which is a very interesting and challenging combination, indeed. Thus, to tie this more general remarks to the foregoing discussion, Williamson is a monist, and Priest is a pluralist. So, they are in stark contrast as far as the issue of *the* right logic is concerned. But they are both anti-exceptionalists. Williamson's position is to defend both classicism *and* monism, and this stance turns on the admission as *the correct logic* of classical logic, and of modal logic, including higher order modal logic. As opposed to that classicist stance, Priest is a dialetheist. I can also offer another example of an exceptionalist who is also a partisan of a non-classical logic. What I mean here is M. Dummett, who is an intuitionist.

But now, it may very well be the case that in order to further clarify the intricacies of the topic, it is worthwhile stress-

[27] Priest 2006a, 2014.

[28] Williamson 2013, 2017a, 2017b.

ing the following point: the difference between exceptionalism and anti-exceptionalism may turn on whether or not logic is normative, and on whether or not logic is empirical or susceptible of being criticized and revised on empirical grounds. But as a general rule, pluralists are not buying into a certain form of exceptional normativism of logic, since they are prone to accept as many logical norms as there are systems of logic; likewise they do not hold that only one logic is the correct logic. End of caveat.

But even this cluster of arguments to the effect that anti-exceptionalism supports logical monism did not go unchallenged recently. And indeed there are philosophers who consider that *anti-exceptionalism* about logic speaks in favor of *logical pluralism* of a more fine-grained sort; namely a kind of *intra-theoretical* pluralism which advocates the idea that different quarters of language obey different sets of logical principles.[29] More to this point, philosophers who speak in favor of logical pluralism upon endorsing an anti-exceptionalist stance about the methodology and epistemology of logic go against a revered tradition and emphasize another dimension of the concept of logic being continuous with that of other more empirical sciences, viz. logic being not topic-neutral, but context-dependent.[30] Henceforth, different domains of knowledge might ask for different logical principles and/or systems, and in order to figure out the logic which fits within the context under scrutiny the best rational strategy is to make use of *abductive reasoning*.[31]

Others, again, have argued for a weaker form of a so-called local logical pluralism on the basis of what they call "full-blooded anti-exceptionalism" about logic.[32] They push to the extreme the idea that the abductive methodology in logic is similar, up to identity, to the methodology of other branches of more empirical sciences. The analogy goes roughly like this: just as "current practice of science" is local and frag-

[29] Hjortland 2017.
[30] Bueno, 2010.
[31] Ibid.
[32] da Costa & Arenhart 2018.

mented, logical inference "comes in variety of kinds that are not as uniform as usually thought".[33]

Despite the many twists and wrinkles of the debate about anti-exceptionalism, and notwithstanding its growing attraction, I would rather say that the whole polemics here is just in its inception. And a lot of difficult issues are in need of clarification; whether or not one should use an abductive methodology in selecting and revising logical systems, which is after all the main pillar of the anti-exceptionalist stance about logic, and, moreover, whether there could exist a genuine compatibility between logical normativity and logical anti-exceptionalism are deep issues which need further careful and thorough investigation.

Where do I stand?

So, after this inventory of various logical "isms", it is quite natural to raise the question, where do I stand? Quite briefly, by and large, I vote for exceptionalism, and with a slight qualification — if not hesitancy — for monism. The metaphysical stance alluded to in the beginning of my paper, which lies in the background, goes like this: there are three distinct sources of necessity — the identity of things, the natural order, the normative order; each produces its own form of necessity; each one is irreducible to the others and there is no more fundamental kind of necessity whose restriction gives rise to those three kinds; any other form of necessity can be understood in terms of either metaphysical, or natural, or normative necessity.[34]

Prima facie, exceptionalism fares better with this pluralistic view about modal and normative concepts. Why is this so? Because it appears to be more in tune and more compatible with this metaphysical outlook according to which normativity cannot be fully naturalized without loss of generality and robustness of those normative thick concepts. Logic captures both the metaphysical and the normative (necessary) truths in a way which can be thought of as being distinct from the ways empirical sciences get to their truths. And this gives logic a quite exceptional status amongst the sciences. Logic, in

[33] Ibid., p. 375.
[34] Fine 2002.

contradistinction to the descriptive empirical sciences, is a *normative* theory/science of rationality.[35] This may mean at least two things:

(i) Logic *does not describe* the psychological processes or states, such as beliefs, desires, thoughts and other propositional attitudes, that are concomitant to reasoning; instead, logic *prescribes* norms, criteria that intervene in the evaluation of the validity or correctness of logical inferences. And the psychological states do not act in a causal manner upon the normative aspects of the correctness of reasoning.

(ii) Logic provides a standard and a set of benchmarks which are instrumental in evaluating whether inferences are correct, or whether they are complying with the standard.

Regardless of the ways we interpret the normativity of logic, all those norms, criteria, canons, standards and so on, and so forth, are *a priori*. Hence, from all those perspectives logic enjoys a kind of exceptional position which gives it a distinctive and exceptional character and makes it singular within the larger family of rational inquiry of the world.

University of Bucharest

References

Biggs, S. & Wilson, J. (2017), "The Apriority of Abduction", *Philosophical Studies* 174, pp. 735–58.

Bueno, O. (2010), "Is Logic A Priori?", *The Harvard Review of Philosophy* 17, pp. 105–117.

Bueno, O. & Colyvan, M. (2004), "Logical Non-Apriorism and the 'Law' of Non-Contradiction", in G. Priest, JC Beall, B. Armour-Garb (2004) (eds.), *The Law of Non-Contradiction*, Oxford University Press, Oxford, pp. 156–175.

Carnap, R. (1937), *The Logical Syntax of Language*, London: Routledge.

Correia F. & B. Schnieder (eds.) (2012), *Metaphysical Grounding: Understanding the Structure of Reality*, Cambridge University Press, Cambridge.

da Costa, N. & Arenhart, J. (2018), "Full-Blooded Anti-Exceptionalism about Logic", *Australasian Journal of Logic* 15, pp. 362–380.

[35] Engel 1991, pp. 306-313.

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

Engel, P. (1991), *The Norm of Truth. An Introduction to the Philosophy of Logic*, University of Toronto Press.

Ferrari, F. (2016), "Disagreement about Taste and Alethic Suberogation", *Philosophical Quarterly*, 66(264), pp. 516-535.

Ferrari, F. (2017), "Logical Pluralism, Indeterminacy and the Normativity of Logic", *Inquiry*, (with S. Moruzzi).

Ferrari, F. (2018a), "Normative Alethic Pluralism", in J. Wyatt et al. (eds.), *Pluralisms in Truth and Logic*, Palgrave MacMillan, London & New York, pp. 145-168.

Ferrari, F. (2018b), "The Value of Minimalist Truth", *Synthese*, 195(3), pp. 1103–1125.

Ferrari, F. (2018c), "Ecumenical Alethic Pluralism", *Canadian Journal of Philosophy*, (with S. Moruzzi).

Field, H. (1996), "The A Prioricity of Logic", *Proceedings of the Aristotelian Society* 96, pp. 359–379.

Field, H. (1998), "Epistemological Nonfactualism and the A Prioricity of Logic", *Philosophical Studies* 92, pp. 1–24.

Field H. (2001), "Apriority as an Evaluative Notion", in H. Field, *Truth and the Absence of Fact*, Oxford University Press, Oxford, pp. 361–391.

Field, H. (2005), "Recent Debates about the A Priori", in T. Gendler & J. Hawthorne (eds.), *Oxford Studies in Epistemology,* Vol. 1, Oxford University Press, Oxford, pp. 69–88.

Field, H. (2009) "What is the Normative Role of Logic?", in *Proceedings of the Aristotelian Society* LXXXIII, pp. 251–268.

Fine, K. (2002), "The Varieties of Necessity", in Fine, K (2005), *Modality and Tense. Philosophical Papers*, Oxford University Press, pp. 235–260.

Fine, K. (2005), *Modality and Tense. Philosophical Papers*, Oxford University Press.

Fine, K. (2012), "A Guide to Ground", in F. Correia & B. Schnieder (2012), pp. 37–80.

Frege, G. (1893), *Grundgesetze der Arithmetik*, English translation by P. Ebert & M. Rosberg with C. Wright, Oxford University Press, Oxford, 2013.

Frege, G. (1956), "The Thought: A Logical Inquiry", in *Mind,* Vol. LXV, No. 259, July, 1956, pp. 289–311.

Hjortland, O.T. (2017), "Anti-Exceptionalism about Logic", *Philosophical Studies* 174, pp. 631–658.

Kant, I. (1781), *Kritik der reinen Vernunft*, English translation by P. Guyer & A. W. Wood, Cambridge University Press, Cambridge, 1998.

Pedersen, N.J.L.L. (2019), "On the Normative Variability of Truth and Logic", *Inquiry*.

Priest, G. (2006a), *In Contradiction* (2nd edition), Oxford University Press, Oxford.

Priest, G. (2006b), *Doubt Truth to be a Liar*, Oxford University Press, Oxford.

Priest, G. (2014), "Revising Logic", in P. Rush (ed.), *The Metaphysics of Logic*, Cambridge University Press, Cambridge, pp. 211–223.

Priest, G. (2016), "Logical Disputes and the a priori", *Logique et Analyse* 59, pp. 347–366.

Putnam, H. (1979), "The Logic of Quantum Physics", *Mathematics, Matter and Method: Philosophical Papers*, vol. 1, 2nd edition, Cambridge University Press, Cambridge, pp. 174–197.

Quine, W.V.O. (1951), "Two Dogmas of Empiricism", *Philosophical Review* 60, pp. 20–43.

Routley, R. (1980), "The Choice of Logical Foundations: Non-classical Choices and the Ultralogic Choice", *Studia Logica* 39, pp. 77–98.

Rumfitt, I. (2015), *The Boundary Stone of Thought*, Oxford University Press, Oxford.

Russell, G. (2015), "The Justification of the Basic Laws of Logic", *Journal of Philosophical Logic* 44, pp. 793–803.

Russell, G. (2017), "Logic isn't normative", *Inquiry*.

Sainsbury, R. (2002), "What Logic Should We Think With?", *Royal Institute of Philosophy* Supplements 51, pp. 1–17.

Sereni, A. & Sforza-Fogliani, M.P. (2017), "How to water a thousand flowers. On the logic of logical pluralism", *Inquiry*.

Steinberger, F. (2017), "Frege and Carnap on the Normativity of Logic", *Synthese* 194, pp. 143–162.

Steinberger, F. (2019), "Three Ways in Which Logic Might Be Normative", *Journal of Philosophy* 116(1),pp. 5-31.

Tarski, A. (1936), "Über den Begriff der logischen Folgerung" in *Actes du Congrès International de Philosophie Scientifique*, fasc. 7 (Actualités Scientifiques et Industrielles, vol. 394), Hermann et Cie, Paris, pp. 1–11.

Tarski, A. (1956), "The Concept of Truth in Formalized Languages", in *A. Tarski: Logic, Semantics, Metamathematics,* Oxford University Press, Oxford, pp. 152–278.

Williamson T. (2013), *Modal Logic as Metaphysics*, Oxford University Press.

Williamson, T. (2017a), "Semantic Paradoxes and Abductive Methodology", in B. Armour-Garb (ed.), *Reflections on the Liar*, Oxford University Press, Oxford, pp. 325–346.

Williamson, T. (2017b), "Model-Building in Philosophy", in R. Blackford and D. Broderick (eds.), *Philosopher's Future: The Problem of Philosophical Progress*, Wiley-Blackwell, Oxford, pp. 159–172.

Wittgenstein, L. (1921), *Tractatus Logico-Philosophicus*, Engl. translation by D. Pears & B. Mc Guinness, Routledge and Kegan Paul, London, 1971.

Woods, J. (2018), "Logical Partisanhood", *Philosophical Studies*.

Wright, C. (1986), "Inventing logical necessity", in J. Butterfield (Ed.), *Language, Mind and Logic*, Cambridge University Press, Cambridge, pp. 187–209.

Are Probabilities Values?

WLODEK RABINOWICZ

According to the Fitting-Attitudes Analysis of value (FA-analysis, for short), for an object to be valuable is for it to be a fitting target of a pro-attitude (i.e., an attitude with a positive valence; in what follows I will often refer to pro-attitudes as 'favorings'). Different kinds of value – desirability, admirability, etc. – correspond to different kinds of fitting pro-attitudes: desire, admiration, etc.

Probability can be interpreted analogously: To be probable is to be credible. Indeed, many probability theorists, from Poisson to Keynes, did adopt this format of analysis.

In earlier publications, I proposed an FA-model of value relations (Rabinowicz 2008, 2012). In this paper, I present a structurally similar model of probability relations. One of the advantages of the model is the interpretation it provides for Keynesian incommensurable probabilities. It goes beyond Keynes, though, in distinguishing between different types of probabilistic incommensurability, some of which Keynes might have been unwilling to allow.

Largely, but not entirely, this paper is a highly condensed summary of the first two parts of Rabinowicz (2017).[1] Among the many changes and additions, the main new element is the argument that credence is a kind of pro-attitude and that probability therefore is a kind of value and not merely a concept that is formally similar to value. The same goes for probability relations: they are value relations of a certain kind.

Another significant new element is the discussion of a distinction between two versions of the FA-analysis of value relations in general and of probability relations in particular.

[1] The material from the third part, on vagueness and probability, has been omitted.

On one form of this analysis, what determines such relations are comparisons between the degrees of fitting pro-attitudes towards different items; for example, we might consider whether it is fitting to favor one item more than the other. On the other version, these relations instead are determined by comparisons of the degrees of fittingness; for example, we might consider whether it is more fitting to favor one item than the other. More fitting to favor is not the same as fitting to favor more. I will argue that the version of FA-analysis that focuses on comparing degrees of fitting pro-attitudes is more plausible than the one that focuses on the comparisons of degrees of fittingness.

1. Values and value relations – FA-analysis

An item is valuable, says the FA-analyst, if and only if it is fitting to favor it. This analysis has an attitudinal and a normative component. While "favor" is a place-holder for a pro-attitude, "fitting" stands for the normative component of the analysis. Other expressions for the normative component might be "warranted", "appropriate", "required", "ought", etc.

An attractive feature of the analysis is that it demystifies value. It makes clear why value judgments are normatively compelling: An object is valuable precisely insofar as we ought to respond to it in a positive way. Normativity is the analytical feature of value. As it stands, the FA-analysis is compatible with competing meta-ethical accounts of its normative component. It is neutral between ethical naturalism, non-naturalism and non-cognitivism.

The analysis makes room for value pluralism: different kinds of value correspond to different kinds of fitting pro-attitudes. Relatedly, it makes room for pluralism regarding value bearers. Objects of different ontological kinds – states, events, actions, persons, concrete things, properties, etc. – can all be targets of fitting pro-attitudes. [2]

FA-analysis readily applies to the standard value relations: betterness and equal goodness. Following Brentano (1969 [1889]) – one of the pioneers of the FA-account – we can as-

[2] For all these features of FA-analysis see Rabinowicz & Rønnow-Rasmussen (2004).

sume that the relevant pro-attitude for these comparative relations is preference. Preference can be understood as a dyadic attitude, directed to two items simultaneously, or, alternatively, as a difference in degree between monadic attitudes of favoring/disfavoring (see Rabinowicz 2012). On the latter interpretation, preference for x over y consists in favoring x to a higher degree than one favors y (or disfavoring x to a lower degree than one disfavors y).[3]

Definitions of the standard value relations are as follows:

> One item is *better* than another iff it is required (fitting, appropriate) to prefer the former to the latter.
>
> Two items are *equally good* iff it is required to equi-prefer them, i.e., to be indifferent.[4]

What if none of the standard relations obtains between two items? To fix terminology, let us say that two items are *incommensurable* in value iff they cannot even be measured on a common ordinal scale, i.e., iff none of them is better nor are they equally good. Ruth Chang has argued that items that are incommensurable in this sense[5] might still be comparable in

---

[3] Such degrees of favoring/disfavoring should however be understood very liberally. They need not be representable by numbers, since the ordering of degrees might not be complete: One degree need not always be higher or lower than another. We have to allow for this, since we want to allow for incomplete preferences (see below).

[4] Two clarifications: (i) Here and in what follows, "iff" stands for "if and only if". (ii) The use of "required" in this analysis might suggest the controversial view that preferences are under our voluntary control. To avoid such commitment to voluntarism regarding pro-attitudes, it might be advisable to instead use normative expressions such as 'fitting' or 'appropriate', which don't have this problematic implication. However, to play the role analogous to the one played in the analysis by "required", these notions should be given a relatively strong reading on which they are equivalent, respectively, to "unfitting not to" and "inappropriate not to." Along with this strong level of normativity, a weak level will be introduced below.

[5] It's not the sense in which she talks of incommensurability though. Chang prefers the more traditional usage: Two items are incommensurable iff they cannot be measured on a common *cardinal* scale (see Chang 2019). This is of course a much less demanding concept of incommensurability than mine.

value: they might be *on a par* (see Chang 1997, 2002a, 2002b, 2005).[6] To use one of her own examples, Mozart and Michelangelo seem to be a case in point: They are on a par in their artistic excellence. As such, they are comparable, even though neither is a better artist than the other. Nor are they equally good: Were we to imagine a slightly improved version of one of them, say, Mozart+ who managed to compose yet another *Requiem* and a few more operas, then this improved version of Mozart would still not be a better artist than Michelangelo.[7] But if Mozart and Michelangelo had been equally good, then anyone better than the former would have to be better than the latter.

As equal goodness, parity is a symmetric relation, but – unlike equal goodness – it is irreflexive (nothing is on a par with itself) and it is not transitive: Mozart+ is on a par with Michelangelo, Michelangelo is on a par with Mozart, but Mozart+ is not on a par with Mozart; he is a better artist.

If parity, as Chang suggests, is a fourth form of comparability in value, along with betterness, worseness and equal goodness, we are left with two notions that need elucidation: parity and comparability.

To define parity, note that the normative component in the FA-analysis can come in a stronger and a weaker form: We can distinguish between requirement and permission, or – in another terminology – between 'ought' and 'may'. The weaker concept is the dual of the stronger one: One may φ iff it is not the case that one ought not to φ. Analogously, it is permissible to φ iff it is not required not to φ.

In Rabinowicz (2008), I suggested this definition of parity:

> x and y are *on a par* iff it is permissible to prefer x to y *and* likewise permissible to prefer y to x.[8]

---

[6] For some notions close to Chang's parity, cf. Griffin's (1986, pp. 81, 96ff) "rough equality", and Parfit's (1984, p. 431) "rough comparability". Cf. also Parfit's (2016) "imprecise equality".

[7] Raz (1986, p. 326) calls the possibility of such a 'one-sided' improvement (or, as the case might be, a one-sided worsening) "the mark of incommensurability".

[8] In case 'requirement' or 'ought' are replaced by expressions that avoid commitment to voluntarism regarding pro-attitudes, corresponding replacements must be made with respect to their duals. Thus, if 'required' is

We might expect to encounter cases of parity when comparisons between items need to take into consideration several dimensions of comparison – several relevant respects. *x* might be superior to *y* in some respects, but inferior in others. Our preference-all-things-considered will then arrive at will depend on the relative weights we give to different dimensions. Different weight-assignments might be admissible and lead to divergent permissible preferences all things considered. When opposing all-things-considered preferences are permissible, we have instances of parity in value.

To prefer one item to another, all things considered, is not the same as to judge it to be better. It is possible to have this preference and still judge both items to be on a par. We are capable of recognizing that the weights we assign to different respects of comparison are not mandatory – that it is admissible to weigh the relevant respects in other ways as well.

To elucidate the notion of comparability, we might ask what *in*comparability would amount to. Since parity is supposed to be a form of comparability and since it at the same time is a typical form of incommensurability, incomparability would have to be an incommensurability of a more radical kind. In Rabinowicz (2008), I made the following suggestion. It seems plausible that preferences all-things-considered might be incomplete. This possibility of 'preference gaps' makes room for incomparabilities within the FA-framework:

> *x* and *y* are *incomparable* iff a preference gap with regard to *x* and *y* is required, i.e., iff if it is impermissible to prefer one of these items to the other or to be indifferent.

Comparability is the negation of incomparability. On this definition, then, parity implies comparability.

If two items belong to different ontological categories, say, one is a state of affairs and the other a person, such items are incomparable in value. It doesn't make sense to prefer one to

---

replaced by 'fitting' or 'appropriate', then 'permissible' has to be replaced by 'not unfitting' and 'not inappropriate', respectively. ("Not unfitting" and "not inappropriate" are duals of "fitting" and "appropriate" if the latter are read as being equivalent with "unfitting not to" and "inappropriate not to", respectively.) I will gloss over this issue in what follows, but it should be kept in mind that such replacements might be advisable.

the other or to equi-prefer them. Similarly, if the items are not the kind of objects to which a given value is applicable, they are incomparable with respect to that value. To illustrate, Mozart and Michelangelo are incomparable in their excellence as scientists: they fall outside the field of that value. But setting such trivial cases aside, it is difficult to find instances of incomparability. A preference gap between items belonging to the same ontological category and falling under the same value might well be permissible in some cases, but it is unclear whether it can ever be required.

2. Value relations – intersection model

What would be an appropriate formal modelling for this approach to different kinds of value relations? An important desideratum is that it should represent 'penumbral connections':[9] While it is not required to prefer Mozart+ to Michelangelo (it is permissible to have the opposite preference), it still is required to prefer the former to the latter if one prefers Mozart to Michelangelo. To impose such 'wide-scope' constraints on combinations of preferential attitudes with regard to different items we need a framework that doesn't deal with each item or each pair of items separately, but instead handles them jointly. This is the main feature of the *intersection model* in Rabinowicz (2008). Its two components are:

> a (non-empty) domain **I** of items under consideration
>
> and
>
> a (non-empty) class **K** of all permissible all-things-considered preference orderings of the items in **I**.

Orderings in **K** need not be complete; preference gaps are allowed. However, in each ordering in **K**, weak preference (i.e., preference-or-indifference) is a quasi-order (or a 'pre-order' as this type of relation also is called): a transitive and reflexive relation. If we take weak preference as a primitive

[9] The idea of penumbral connections is due to Kit Fine. It was introduced in his seminal paper on the supervaluationist modelling of vagueness (Fine 1975). I use it here in a different, but closely related sense. The intersection modeling to be presented below is in some respects formally similar to the supervaluationist model.

concept, then preference and indifference can be defined as, respectively, the asymmetric and the symmetric parts of weak preference. This entails, in particular, that both permissible preference and permissible indifference are transitive.[10]

The relation of betterness between items can be defined as the intersection – the common part – of permissible preferences.

> *x* is *better* than *y* iff *x* is preferred to *y* in every ordering in **K**.

Since what holds in all permissible orderings is *ipso facto* required, this is another way of saying that *x* is better than *y* iff preferring *x* to *y* is required.

Other value relations are defined correspondingly:

> *x* and *y* are e*qually good* iff they are equi-preferred in every **K**-ordering;
>
> *x* and *y* are *incomparable* iff there is a gap between them in every **K**-ordering;
>
> *x* and *y* are *on a par* iff *x* is preferred to *y* in some **K**-orderings and *y* is preferred to *x* in some other **K**-orderings;
>
> and so on.

The intersection model has no difficulties with constraints on the combinations of preferential attitudes. In the Mozart-Michelangelo example, every **K**-ordering would place Mozart+ above Mozart, which means that it would also place Mozart+ above Michelangelo if it places Mozart above Michelangelo. In this sense, then, there is a wide-scope con-

---

[10] The elements of **K** might be thought of as permissible "preferential states" of a person. Instead of representing a preferential state as I have done, as a partial ordering, we could instead represent it as a set of complete orderings such that their intersection coincides with the partial ordering under consideration. A person who would be in such a state might be thought of as being undecided between the different complete preference orderings in the set. This representation is more fine-grained than the one I have chosen, since different sets of complete orderings can have the same intersection. But we probably don't need this level of grain. For our purposes, representing preferential states as partial orderings should be sufficient.

straint on preferences: It is required that (one prefers Mozart+ to Michelangelo if one prefers Mozart to Michelangelo).

The intersection model allows us to derive formal properties of value relations from the constraints on permissible preference orderings. Given that every permissible weak preference is a quasi-order, it follows that betterness is a transitive and asymmetric relation, that equal goodness is an equivalence relation, and that whatever is better than, worse than, on a par with, or incomparable with one of two equally good items must stand exactly the same value relation to the other item.[11]

To put the above more succinctly, it follows in our model that, just as permissible weak preference, "better than or equally as good as" is a quasi-order: a relation that is reflexive and transitive, but not necessarily complete. Incommensurabilities in value are allowed.[12]

We now also have all we need for a general classification of binary value relations:

---

[11] The model also makes it clear what needs to be done if one, like Temkin (2012), wants to reject the transitivity of betterness: One needs to give up the corresponding transitivity constraint on permissible preference orderings. But note that, on the interpretation of preference as favoring more or disfavoring less, transitivity is an analytic feature of preference in general and not merely a constraint of permissible preferences, which may or may not be accepted (see Rabinowicz 2012). Consequently, denying that betterness is transitive presupposes, at a minimum, that preference is interpreted as a dyadic attitude and not as a difference in degree between monadic attitudes of favoring/disfavoring.

[12] Szpilrajn's extension theorem states that a strict partial order (an irreflexive and transitive relation) can always be extended to a total order (to a linear order, in other words). For infinite domains, the proof of this result relies on Zorn's lemma. An immediate corollary is that a quasi-order can always be extended to a weak order: a quasi-order without gaps. So why do we insist that the "at least as good as" relation might have to contain gaps (incommensurabilities) if these gaps can always be filled in by extending this relation to a weak order? (I am indebted to an anonymous referee for pressing this point.) Well, the answer is that a gappy quasi-order has many different non-gappy extensions and therefore choosing one of them rather than one of the others would be wholly arbitrary. Indeed, this applies to any gap in the quasi-order. For any particular way of filling it in no justification can be provided. We have to live with incommensurabilities in value.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\succ$	+			+		+	+	+	+	+			+		
$\prec$		+			+	+	+	+	+		+			+	
$\approx$			+	+	+		+	+				+	+	+	
/								+	+	+	+	+	+	+	+
	B	**W**	**E**			**P**	**P**	**P**	**P**						**I**

In the table above, each column corresponds to one type of a value relation that might obtain between two items. In each column, the preferential stances that in this type are permissible with regard to an ordered pair of items are marked with plus signs. There are four such stances, corresponding to the rows in the table: preference ($\succ$), its converse – dispreference ($\prec$), indifference ($\approx$), and a gap (/). Since with regard to each ordered pair of items at least one preferential stance must be permissible, there are fifteen possible ways of picking out permissible stances: 2^4-1 = 15. Consequently, there are fifteen columns in the table – fifteen types of value relations. In types 4 - 14, several preferential stances are permissible, while in types 1, 2, 3 and 15, which correspond to betterness (**B**), worseness (**W**, the converse of betterness), equal goodness (**E**) and incomparability (**I**), the only permissible stances are, respectively, preference, dispreference, indifference and a gap.

The table's columns stand for *atomic* types of value relations. Parity (**P**) is not an atomic type; it is a collection of types 6 - 9, i.e., the types in which there are plus signs in both the first and the second row. Such collections of atomic types, as parity and comparability (types 1 - 14), form types in a broader sense. In this respect, parity differs from the three traditional value relations: **B**, **W** and **E**.

The fifteen atomic types are all *logically* possible, but some of them might not represent 'real' possibilities: they might not be instantiated in the domain under consideration. Thus, for example, incomparability might not be such a real possibility. Likewise, one might reject as real possibilities some of the parity types. It is arguable that both indifference and preferential gaps should always be permissible with regard to items that are on a par. This would exclude types 6, 7 and 9. In the

space of real possibilities parity might thus boil down to one atomic type: type 8.

3. Probabilities and probability relations – FA-analysis

The suggestion I now want to make is that our ordinary notion of probability lends itself to FA-analysis: *To be probable is to be credible*. A proposition is probable in this sense iff it is fitting to give it credence. This makes probability a kind of value *if* credence, or more precisely high credence, is a kind of pro-attitude. Indeed, it does seem that high credence – belief – is a pro-attitude of sorts, while low credence – disbelief – is the corresponding con-attitude. The former is a form of favoring, the latter is a form of disfavoring. The language we use when we talk about doxastic attitudes is strongly suggestive in this regard. We *accept* some views and *reject* others. We *embrace* what we take to be true and *spurn* what we take to be false.

One might object that beliefs and favorings, such as desires, have opposite 'directions of fit': mind-to-[fit]-world and world-to-[fit]-mind, respectively. We adjust our beliefs to what is the case and adjust what is the case to our desires.[13] But this objection is not convincing. There are other typical forms of favoring, such as admiration, awe, approbation or joy, that have the same direction of fit as belief. In admiration or approbation I am adjusting my mind to the world, just as in the case of belief.

But, the objector might persist, what makes admiration or approbation pro-attitudes is that they contain some conative, desire-like component, even if they primarily have the mind-

[13] The distinction between two directions of fit is often ascribed to Anscombe 1957, even though she never made it in these terms. For a good discussion of direction of fit, see Humberstone (1992). Actually, the idea of two directions of fit is much older. We can find it already in Aquinas's *Summa Theologica* (part 1, question 21, article 2) as a distinction between two forms of truth. For Aquinas truth consisted in a fit between things and mind (*adaequatio rei et intellectus*). He suggested that such fit can be realized in two ways: Either mind fits itself to things (as when it "receives knowledge from things") or it fits things to itself (as when an artist creates a work of art "in accordance with his art", or as when "God's justice" "establishes things in the order conformable to the rule of His wisdom").

to-world direction of fit. In admiring a person, I might, or perhaps even must, desire to be like her, in approving of a state of affairs, I desire it to continue, etc. Even this variant of the objection can be dealt with, I think. A conative component is present in belief as well: Belief is a form of trust. To trust x is to be willing to rely on x. I am willing to rely on what I believe – to act on this proposition. This willingness to rely is a conative attitude that is intimately related to believing.[14]

Thus, it does seem that probability may well be viewed as a value. It might be noted, by the way, that in the classical Latin "*probabilis*" was often used to mean "worthy of approval, pleasing, agreeable, acceptable, commendable, laudable, good, fit" (Lewis & Short 1879). That "*probabilis*" was a term of evaluation must have been evident to the Romans.

As applied to propositions to which we assign quantitative degrees of probability, the FA-style account of probability I suggest we adopt takes the following form:

A proposition A is *probable to degree k* iff A ought to be given credence of degree k.[15]

Here and in what follows, I assume that the relevant 'ought' is meant to be relativized to available evidence. It is the latter that determines what credence one ought to assign to a proposition under consideration.[16]

---

[14] Indeed, on the influential betting interpretation of credences, an agent's degree of credence in a given proposition is accounted for, or at least measured by, her willingness to bet on the proposition in question.

[15] A bi-conditional like this might be accepted even by someone who isn't prepared to view it as an *analysis* of probability but merely as an adequacy condition on any satisfactory analysis of this notion. The modelling I develop in what follows is compatible with such a cautious position.

[16] There is a problem here that is a special case of a general difficulty facing FA-analysis: Some of the reasons why one ought to give high or low credence to a proposition, A, might be of the "wrong kind": they might be unrelated to what makes A more or less probable. In the FA-analysis of value, defining the distinction between reasons for favoring that are of the wrong kind and those of the right kind seems difficult, if not impossible, without already assuming the notion of value that is being analyzed (see Rabinowicz & Rønnow-Rasmussen 2004). When it comes to probabilities, reasons of the right kind are supposed to be evidentiary (in contrast to, say, pragmatic or moral reasons for belief). But it is unclear whether we can define what we mean by evidentiary reasons in a non-circular way,

The FA-analysis of probability contains an *attitudinal* component (credence) and a *normative* component (ought). It thus makes probability an explicitly normative notion. In this respect, it differs from subjective probability accounts, which simply identify probability with credence. To the extent the subjective accounts are normative, their normativity is implicit and purely formal or structural: It comes to expression in various formal constraints that credences are supposed to satisfy (Kolmogorov's axioms, the Reflection Principle, conditionalization, etc.). By contrast, the FA-account wears normativity on its sleeve. And its normativity is not restricted to purely formal constraints. The evidence available to a subject will typically impose substantive constraints on her credences.[17]

As compared with the subjective accounts, the FA-analysis much better squares with our ordinary understanding of probability. On that understanding, probability judgments are meant to guide our credences rather than to express or report them. This guiding function of probability judgments immediately follows given the FA-account. A closely related consideration is that, on the ordinary understanding, an agent's credences might sometimes come into conflict with her probability judgments. This is likewise in agreement with the FA-account but not with the subjective accounts, at least not without making them considerably more complicated.

The normative character of the FA-account also makes it attractive when we compare it with theories of 'objective' probability (or 'chance', as it often is called), such as fre-

---

without bringing probability at some point. We might want to say that such reasons support the proposition under consideration or its negation. But can this notion of support be clarified without appealing to the concept of probability? Can we clarify it in some other way than in terms of increase in probability? If not, circularity threatens. I gloss over this problem in what follows.

[17] On this account, the normativity of probability is that of epistemic rationality. It is clearly not the normativity of morality; nor is it the normativity of prudence, at least not primarily so (though, as we have noted above, credences involve dispositions to action). Does it mean that different kinds of values aren't normative in the same way? Or just that their normativity has different sources? I am not sure. I am indebted to an anonymous referee for raising this issue.

quentism or the propensity theory. On the objectivist accounts, the role of probabilities as a guide to credences is not explicit; it needs to be argued for. The Principal Principle, according to which credences ought to reflect expected chances, is a substantive claim. But no argument is needed to establish that credences ought to track probabilities, if the latter are understood in accordance with the FA-account. It should be noted, however, that accepting the FA-account does not in any way require rejecting the concept of objective probability. There is ample room for that concept as well, as it might still be needed in our descriptions of the world.[18]

The FA-account of probability is sometimes referred to as "the epistemic interpretation" of this concept.[19] Just as the general FA-account of value, this approach to probability has a history. Among its early proponents we find nineteenth century probability theorists such as Poisson (1837, p. 31), Cournot (1843, pp. V, 438), de Morgan (1847, pp. 172f), Donkin (1851, pp. 355f), Boole (2009 [1854], pp. 187f) and Jevons (1873, p. 198).

Poisson seems to have been the pioneer in this development. It is misleading, though, to count him as the representative of the account of probability I am interested in. He takes an event to be more probable on one kind of evidence than on another if the former evidence gives us "more grounds to believe" it.[20] He also identifies what is more prob-

---

[18] It might be possible, though, to analyze objective probabilities in terms of FA-probabilities, on the lines of David Lewis's (1980) well-known reduction of objective chances to credences that would be rational (i.e., fitting) given appropriate hypothetical restrictions on available evidence.

[19] But the terminology varies. In Galavotti (2011), "the epistemic interpretations" are subdivided into two categories: the "subjective" and the "logical" interpretations. The adherents of the FA-approach are included in the latter category, which culminates with Carnap. For another useful recent historical account, see Zabell (2011).

[20] "The *probability* of an event is our reason to believe that it will occur or occurred. […] Probability depends on our knowledge about an event; for the same event it can differ for different persons. Thus, if a person only knows that an urn contains white and black balls, whereas another person also knows that there are more white balls than black ones, the latter has more grounds to believe in the extraction of a white ball. In other words,

able with what there is "a greater reason to believe", or what we have "more reason to believe" (Poisson, p. 32). But it is not the same as reason to believe, to a higher degree, as the analysis I am interested in would have it. Indeed, more reason to believe, or stronger reasons to believe, are not even necessarily co-extensive with reasons for stronger belief.

It might be of interest to note that an account of probability similar to Poisson's has recently been defended by John Skorupski:

> Talk of probability is talk about epistemic reasons. 'It's probable that *p*' means 'There's reason of some degree to believe that *p*'. 'There's an *m*/*n* probability that *p*' means 'There's reason of degree *m*/*n* to believe that *p*'. This normative definition, we shall argue, is correct. (Skorupski 2010, p. 220.)

Can one grade reason to believe in this way, on a scale from 0 to 1? Isn't it quite counter-intuitive? And why does Skorupski want to grade the strength of reason to believe rather than the strength of belief there is reason to have?

The answer comes later in the same chapter of his book (ibid., ch. 9):

> The idea of a degree of belief is obscure. Belief, unlike feeling, does not come in degrees. One can feel more or less strongly, but one cannot believe more or less strongly. […] True, we can be more or less sure; more or less *inclined* or *disposed* to believe. (Ibid., p. 228; see also p. 51.)

In my opinion, Skorupski's claim that belief doesn't come in degrees is simply wrong, though it is not something I can argue for here.[21] He is in any case perfectly aware that this

---

for him, that event has a higher probability than for the former." (Poisson, ibid.)

[21] Beliefs are intimately connected with dispositions to act. Therefore, the standard way of measuring degrees of belief is by exposing the subject to an array of bet offers. This idea, most famously defended by Ramsey and de Finetti, can be traced back at least to Kant's *Kritik der reinen Vernunft* (A824-5, B852-3). While in my view betting dispositions do provide a good measure of the strength of belief, this measure is not without its problems. One well-known problem is that one cannot make bets on undecidable propositions, another is that betting dispositions also depend on one's attitudes to risk. Also, there is an interference problem: In some

claim is in stark conflict with the nowadays dominant Bayesian accounts of belief.

In value theory, a proposal on the same lines as Poisson's came more than a half century later. It was brought forward by G. E. Moore (1903), in his critical review of Brentano (1969 [1889]). Moore's suggestion was that, if 'good' is analyzed as what is 'worthy of love', as Brentano took it to be, then 'better' (i.e., 'more good') should not be analyzed, as Brentano suggested, as what is worthy of greater love, but as what is more worthy of love.[22] To put it differently and more generally, on this 'Moorean' kind of view, x is more valuable than y iff it is more fitting to favor x than y, i.e., iff there are stronger reasons to favor x than to favor y.

I don't find this Moorean version of the FA-analysis plausible. To argue against it would require an extended discussion, but let me just point out two unwelcome implications of his approach. The basic problem is that this approach does not give us any guidance as to the degree of a pro-attitude we ought to have towards a valuable object, even if the degree of its value is specified. If an object is highly desirable, ought it to be highly desired? If a proposition is not very probable, ought it to be given a low credence? These questions are left hanging and they shouldn't be. On the other hand, such an approach gives guidance where it shouldn't. It makes comparisons between degrees of different values easier than they intuitively ought to be. Think of statements such as this one: "x is more admirable than y is desirable." On the Moorean approach, such a comparison would be relatively unproblematic, since there is a common currency in terms of which it can be made. It would mean that x is more worthy of admiration than y is worthy of desire. Or, to put it in terms of the

---

cases, the very offer of a bet on a proposition can affect one's degree of belief in that proposition (see Eriksson & Rabinowicz 2012). But that a measure has its faults does not show, of course, that what is being measured does not exist.

[22] "[Brentano's] first suggestion is that since 'good' means 'worthy to be loved,' 'better' must mean 'worthy of more love' (p. 21). It does not seem to have occurred to him that it must mean 'more worthy of love,' that is to say, his attention is directed only to that element of his definition, which is a 'concrete psychical content,' namely the love, not to the more important element 'rightness,' which is not." (Moore 1902, p. 118.)

strength of reasons, it would mean that there are stronger reasons to admire *x* than to desire *y*. But comparisons of degrees of admirability with degrees of desirability should *not* be unproblematic. It is not even clear whether such cross-categorial comparisons make sense.

In the same way, Poisson's analysis of probability is questionable. Indeed, in this case, as we have seen in connection with Skorupski's proposal, there is an additional issue of whether reasons can be meaningfully graded in the way probabilities can, on a scale from 0 to 1. But already de Morgan (1847) explicitly defined probability as the degree of belief one ought to have, given evidence. What is graded on de Morgan's account is the belief one has reasons to have and not the reasons for this belief – not, as we might say, the weight of evidence in favor of this belief. It is a more satisfactory version of the FA-analysis of probability than Poisson's; the probability of a proposition is not the same as the weight of available evidence that supports a given degree of belief in the proposition in question. When the latter increases, the probability of the proposition does not change, but it becomes more resilient to potential later changes when further evidence comes in.[23]

---

[23] Indeed, the weight of evidence is not, as such, even a tie-breaker when it comes to probability comparisons. Suppose one ought to give a certain credence to a proposition *A* and the same credence to a proposition *B*, but the reasons for giving this credence to *A* are stronger. For example, let *A* be the proposition that coin *a* will land heads at least once in the next two tosses, while *B* is the same proposition about another coin, *b*. Both *a* and *b* have been tested and no asymmetries in their shapes and distributions of weight have been found. But the tests of *a* have been more thorough. Is *A* then more probable than *B*, given this evidence? I would deny it. If the two propositions ought to be given the same degree of credence, then they are equally credible – equally probable. That the weight of evidence is greater with regard to *A* than with regard to B doesn't change this, though it makes the probability of *A* more resilient: less sensitive to further evidence.

The same applies to values in general. The strength of reasons is not even a tie-breaker when it comes to value comparisons. Thus, for example, suppose that one ought to desire *x* in a certain degree and to desire *y* in the same degree, but the reasons to desire *x* in this degree are stronger. It wouldn't be plausible to draw the conclusion that *x* is more desirable

It was de Morgan's account that subsequently became dominant. This development culminated with J. M. Keynes's *A Treatise on Probability* (1921). Here are some representative quotes from the first pages of this book:

> The terms *certain* and *probable* describe the various degrees of rational belief about a proposition which different amounts of knowledge authorise us to entertain. (p. 2)

> The Theory of Probability [...] is concerned with the degree of belief which it is *rational* to entertain in given conditions, and not merely with the actual beliefs of particular individuals, which may or may not be rational. (p. 3)

> A *definition* of probability is not possible, unless it contents us to define degrees of the probability relation by reference to degrees of rational belief. We cannot analyse the probability-relation in terms of simpler ideas. (p. 7)

"Rational" stands for the normative component in Keynes's analysis and "knowledge" stands for evidence.

Keynes also explicitly distinguishes the probability of a proposition – the degree of belief that it calls for – from the weight of evidence concerning the proposition in question. He often refers to the latter as "the weight of an argument":

> As the relevant evidence at our disposal increases, the magnitude of the probability of the argument may either decrease or increase, according as the new knowledge strengthens the unfavorable or the favorable evidence; but something seems to have increased in either case,—we have a more substantial basis upon which to rest our conclusion. I express this by saying that an accession of new evidence increases the weight of an argument.

---

than *y*. If the degree of desire ought to be the same with regard to both, then they are equally desirable.

A more reasonable position is to view the greater strength of reasons not as a factor that increases the value of the object under consideration, but instead as the determinant of the resilience of value to potential changes in the underlying reasons, just as in the case or probabilities. (Cf. Andersson & Green Werkmäster 2020.)

> New evidence will sometimes decrease the probability of an argument, but it will always increase its 'weight.' (Ibid., p. 78.)[24,25]

Keynes combines his FA-account of probability with strict objectivism concerning what is rational to believe given available evidence:

> When once the facts are given which determine our knowledge, what is probable or improbable in these circumstances has been fixed objectively, and is independent of our opinion. (p. 3)

But FA-analysis does not by itself presuppose any such objectivism. As I have mentioned above, in connection with the analysis of value, the FA-account is compatible with different interpretations of its normative component, including subjectivism and even non-cognitivism.

The history of the FA-account of probability does not end with Keynes. Among its later proponents we find such philosophers as as Jeffreys (1931, 1939) and Carnap (1962,

---

[24] It is not clear to me whether Keynes is right in this general claim. Cannot new evidence sometimes decrease "the weight of an argument"? What about new evidence to the effect that old evidence for proposition *A* is not to be relied on? Say, we might learn that a witness previously thought to be reliable is not to be trusted. The testimony he has given in support of *A* is thereby undermined. In a case like this, the weight of evidence concerning *A* seems to be decreased by new information.

[25] While the greater weight of evidence does not, as such, increase the probability of a proposition, Keynes takes it to be an important factor in decision making. In ch. XXVI, dealing with "application of probability to conduct", he writes: "the degree of completeness of the information upon which a probability is based does seem to be relevant, as well as the actual magnitude of the probability, in making practical decisions. […] If, for one alternative, the available information is necessarily small, that does not seem to be a consideration which ought to be left out of account altogether." (pp. 357f) This is not the only deviation on Keynes's part from the standard principle of the maximization of expected value. He also suggests that the marginal contribution of value increases is decreasing and that, as a result, "a smaller but relatively certain good is better than a greater but proportionately more uncertain good", even if the expected goodness is the same in both cases (p. 366). He develops both of these suggestions in ch. XXVI.

1971).[26] But it is Keynes who will play the central role in my discussion below, because of his seminal role when it comes to incommensurable probabilities.

The FA-account of probability lends itself to the treatment of probability relations that is exactly analogous to our treatment of value relations in general. Indeed, if probability is a kind of value, it is more than an analogy. Just as value relations are defined in terms of preference, probability relations are defined in terms of differences in credence. If high credence is a form of favoring and low credence the corresponding form of disfavoring, then to give more credence is to favor more or to disfavor less. Consequently, on the interpretation of preference as a comparative relation between monadic attitudes of favoring/disfavoring, giving more credence is a way of preferring.

Letting *A* and *B* vary over propositions, we get the following series of definitions:

> *A* is *more probable* than *B* iff *A* ought to be given more credence than *B*.
>
> *A* and *B* are *equiprobable* iff *A* and *B* ought to be given equal credence.
>
> *A* and *B* are *incommensurable* (probabilitywise) iff neither is more probable than the other nor are they equiprobable.
>
> *A* and *B* are *on a par* (probabilitywise) iff it is permissible to give *A* more credence to than *B* and permissible to give *B* more credence than *A*.
>
> *A* and *B* are *incomparable* (probabilitywise) iff if it is impermissible to give either more credence than the other or to give both equal credence.

---

[26] In his (1971, p. 8), Carnap stresses the need to distinguish between "two versions of personal probability, one representing the *actual* degree of belief and the other the *rational* degree of belief." Rational degrees of belief are taken to satisfy standard axioms of probability and they also are appropriately constrained by the person's evidence. His conception of logical (or, as he sometimes calls it, "inductive" probability) is a further and much more radical elaboration of this concept of rational degrees of belief.

What is required in case of incomparability is a *credence gap*. In case of such gap, none of the propositions is given more credence than the other nor are they given equal credence. How to model credence gaps will be considered later. As before, *comparability* is defined as the negation of incomparability.

Can different types of probabilistic incommensurability be instantiated in real life? As for probabilistic parity, we might expect to encounter it when different parts of our evidence come from different sources or are of different kinds. To assign credences to propositions for which this evidence is relevant, we will need to weigh its parts against each other. Different assignments of weights might be admissible and give rise to parity. If one part of evidence supports *A* while another supports *B*, then it might well be permissible to give higher credence to *A* than to *B* and likewise permissible to give higher credence to *B* than to *A*. In such a situation, we might (permissibly) give more credence to *A* than to *B* and yet deny that that *A* is more probable. Just as we might (permissibly) prefer one item to another and yet deny that it is better.

But, and this should also be stressed, in some cases it might turn out that there is *no* admissible way of weighing different parts of mixed evidence, or – even if there are such ways – it is at least permissible to abstain from weighing.

Keynes describes the following case:

> Is our expectation of rain, when we start out for a walk, always more likely than not, or less likely than not, or as likely as not? I am prepared to argue that on some occasions none of these alternatives hold, and that it will be an arbitrary matter to decide for or against the umbrella. If the barometer is high, but the clouds are black, it is not always rational that one should prevail over the other in our minds, *or even that we should balance them*,— though it will be rational to allow caprice to determine us and to waste no time on the debate. (Keynes 1921, pp. 31f, my italics.)

Keynes is saying that in a case like this, with different parts of evidence ("the barometer is high, but the clouds are black") supporting opposing hypotheses (No Rain and Rain, respectively), it might be rational to make an arbitrary choice of action ("to decide for or against the umbrella", "to allow caprice

to determine us"), without weighing different parts of the evidence against each other. It might not be rational to give one part of the evidence more weight than the other, or to give them equal weight ("balance them"). It might therefore not be rational to give one hypothesis more credence than the other or give them equal credence. But does "not rational" mean "irrational" here, or just "not rationally required"? Whether a credence gap is required in such a case or merely permissible is not clear from this passage. I will come back to this question in the last section.

4. Intersection model for probabilities

Moving now to formal representation, how can the different probability relations be modelled? We need something structurally similar to the *intersection* model I presented in section 2 above. In this case, the domain **I** of items consists of propositions – the objects of credence. **I** is non-empty and closed under Boolean operations. For simplicity, I take it that **I** is finite, but extension to the infinite case does not pose special problems. The class **K** is now the non-empty set of all permissible *doxastic states*. Intuitively, a doxastic state *S* is a possible credential state of a person. It is permissible if it is permissible to be in such a state given available evidence. (The latter is held fixed here.) We represent a permissible state *S* formally as a non-empty set of credence functions, with **I** as their common domain. Each credence function assigns a numerical value, between 0 and 1, to every proposition in the domain and it satisfies Kolmogorov axioms.

A state *S* is *fully opinionated* if it is a singleton. A person in that state has a definite credence assignment to every proposition. *S* is not fully opinionated if it contains several credence functions. In such a state, a person is undecided to some extent, when it comes to her credences. This way of representing doxastic states is quite standard in the literature on

indeterminate (imprecise or ambiguous[27]) credences. In philosophy, it goes back to Levi (1980).[28]

We can now define what it means that in a doxastic state, a proposition is given more credence than another proposition or that two propositions are given equal credence. We can also define the notion of a credence gap.

> *A* is given *more credence* than *B* in *S* iff, for every credence function *C* in *S*, $C(A) > C(B)$.
>
> *A* and *B* are given *equal credence* in *S* iff, for every *C* in *S*, $C(A) = C(B)$.
>
> There is a *credence gap* with regard to *A* and *B* in *S* iff none of *A* and *B* is give more credence in *S*, nor are they given equal credence in *S*.

Thus, there is a credence gap in *S* with regard to *A* and *B* iff the credence functions in *S* do not rank these two propositions uniformly vis-à-vis each other. Credence gaps can only arise in doxastic states that are not fully opinionated.

We can now also interpret various quantitative claims about doxastic states. For example, *A* is given credence *k* in *S* iff, for every *C* in *S*, $C(A) = k$. *A* is given twice as high credence as *B* in *S* iff, for every *C* in *S*, $C(A) = 2C(B)$. And so on.

There is an important formal dissimilarity between my intersection model for values in general (section 2) and this model for probabilities. In the former, permissible preferential states are represented as partial orderings of items in the domain. Why are doxastic states not represented in the same purely qualitative way, as partial credence orderings of propositions? The answer is that in the case of probabilities we want to model not just qualitative probability relations, but also quantitative claims about probabilities, to the extent that such claims can be made. If we had the same ambitions for

---

[27] Imprecision and ambiguity of credences are not quite the same phenomena. It is ambiguity that has more to do with indecision – with absence of commitment, while imprecision is more related to the definiteness of one's commitments. But here I gloss over this issue.

[28] For a succinct discussion of the advantages of this representation and for some further information on its historical roots in economics and statistics, see Weatherson (2002). Cf. also Weatherson (2007).

the value model, we would need to do something similar in that case as well. For example, if we think that values can be to some extent cardinally compared, we can represent preferential states as sets of cardinal preferences, i.e., as sets of cardinal measures of the items in the domain (with each such measure being in turn representable by a set of numerical functions closed under affine transformations).

Setting aside, for the time being, quantitative statements about probabilities, we now have what we need to define qualitative probability relations in terms of the intersection model. The definitions follow, of course, the format of the definitions of value relations in general.

> *A* is *more probable* than *B* iff *A* is given more credence than *B* in every *S* in **K**.
>
> *A* and *B* are *equiprobable* iff *A* and *B* are given equal credence in every *S* in **K**.
>
> *A* and *B* are *on a par* (probabilitywise) iff *A* is given more credence than *B* in some *S* in **K** and *B* is given more credence than *A* in some other *S* in **K**.
>
> *A* and *B* are *incomparable* (probabilitywise) iff there is a credence gap with regard to *A* and *B* in every *S* in **K**.

And so on. We can also provide a classification of binary probability relations that is exactly analogous to the corresponding classification for value relations in general:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
≻	+			+		+	+	+	+	+			+		
≺		+			+	+	+	+	+		+			+	
≈			+	+	+		+	+				+	+	+	
/								+	+	+	+	+	+	+	+
	M	**L**	**E**			**P**	**P**	**P**	**P**						**I**

Here, the four rows correspond to four possible credential stances with regard to an ordered pair of propositions: giving more credence to the first proposition in the pair (≻), giving it less credence (≺), giving both propositions equal credence (≈), and being in a state of credence gap (/). Again, there are

fifteen columns, standing for the different atomic types of probability relations, with each type being defined by the credential stances that are permissible in the type in question. The permissible stances are marked with the plus signs in the column. As for the type labels, we still have **E**, **I** and **P**, which now stand for equal probability, probabilistic incomparability and probababilistic parity, respectively. (**P** is again a collection of four atomic types.) But **B** (better) and **W** (worse) are now replaced by **M** (more probable) and **L** (less probable).

While the fifteen atomic types are all logically possible, some might lack real instantiations. If we are after 'real possibilities', we might want to reduce the table by two kinds of restrictions: (i) constraints on every doxastic state in **K**, taken separately, or (ii) constraints on interrelations between doxastic states, i.e., restrictions on class **K** taken as a whole.

An example of a restriction of the first kind is the condition stating that all permissible doxatic states are fully opinionated. This would imply that no credence gaps are permissible and thus remove all types from 8 to 15.

As for restrictions of the second kind, we might require that in all cases of probabilistic parity both equal credence and a credence gap should be permissible. This would remove types 6, 7 and 9, and reduce parity to type 8.

Several other restrictions of the second kind might be worthy of consideration, but the following very radical restriction especially deserves attention:

Uniqueness: There is only one permissible doxastic state.

Given Uniqueness, there is no real difference between what is required (what holds in all permissible states) and what is permissible (what holds in some permissible state). Consequently, for any pair of propositions only one credential stance is permissible with regard to that pair. Therefore, all types apart from the first three and the last one are removed. The last type – incomparability – is instantiated if the unique permissible doxastic state contains credence gaps. Uniqueness implies that probabilistic incommensurability boils down to incomparability.

Uniqueness will be an appealing condition for an adherent of the FA-account who is unwilling to entertain probabilistic parity – unwilling to allow that, for some A and B, it may be

permissible to give more credence to *A* than to *B* and likewise permissible to give more credence to *B* than to *A*. If the evidence does not require giving more credence to *A* nor does it require giving more credence to *B*, and if it in addition does not require giving *A* and *B* equal credence, then – on this view – the only permissible stance open to the agent is to maintain a credence gap with regard to the propositions in question. I will have more to say about Uniqueness in the last section.

Constraints mentioned above are just some examples of restrictions that might be imposed on **K**. But all these extra constraints go beyond the basic intersection model and, as such, should be treated with caution: They should not be imposed unless we are confident that they hold.

5. Keynes

One of the goals of the paper has been to provide a framework that accounts for incommensurable probabilities, along the lines of Keynes's *Treatise* (1921). The objective has been to make room for propositions whose probabilities cannot be specified numerically. Note that such non-quantifiable probabilities might sometimes enter into quantitative relations with each other. For example, it is in principle possible in our model to have two propositions such that the ratio between their probabilities is fixed at a definite number, say, one might be twice as probable as the other, even though these propositions cannot be assigned numerical probabilities. It is enough if the ratio between their credences is constant across all credence functions in all permissible doxastic states. In other cases, it will be possible to measure the probability difference between two propositions with non-quantifiable probabilities – if the difference between their credences is the same in all credence functions in all permissible states.

The intersection model allows assigning numerical probability values to some propositions. This will be the case for any proposition whose credence is constant across all credence functions in all doxastic states in **K**. But if **K** is large or if some states in **K** contain many credence functions, this will be an exception rather than a rule. Likewise, it will be an exception rather than as a rule that the probabilities of different

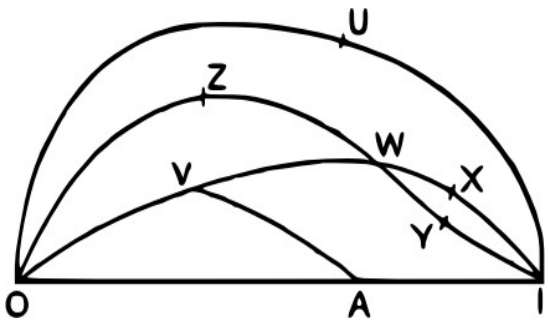
propositions can be quantitatively compared. And in some cases they won't be even commensurable, i.e., ordinally comparable.

This is in line with Keynes's ideas:

> I maintain [...] that there are some pairs of probabilities between the members of which no comparison of magnitude is possible […] and that in a very special type of case [...] meaning can be given to a numerical comparison of magnitude. […]
>
> By saying that not all probabilities are measurable, I mean that it is not possible to say of every pair of conclusions, about which we have some knowledge, that the degree of our rational belief in one bears any numerical relation to the degree of our rational belief in the other; and by saying that not all probabilities are comparable in respect of more and less, I mean that it is not always possible to say that the degree of our rational belief in one conclusion is either equal to, greater than, or less than the degree of our belief in another. (Keynes 1921, p. 36)
>
> There are, however, many cases […] in which it is certain that it is not actually within our power to make the comparison. It has been argued that in these cases the probabilities are, in fact, *not comparable*. As in the example of similarity, where there are different orders of increasing and diminishing similarity, but where it is not possible to say of every pair of objects which of them is on the whole the more like a third object, so there are different orders of probability, and probabilities, which are not of the same order, cannot be compared." (Ibid., p. 124.)[29]

The intersection model bears out Keynes's diagram representing probability space (ibid., p. 42):

[29] How important the idea of incommensurabilities was for Keynes (especially, but not exclusively, in his earlier work), not just in theory of probability but in other areas as well, ranging from ethics to economics, is well described in Carabelli (1998).



O stands for "impossibility" (the impossible proposition, assigned probability 0), I for "certainty (the certain proposition, assigned probability 1). Propositions on the same path are linearly ordered from from left to right by increasing probability (example: W is more probable than V). Such paths are what Keynes calls "orders of probability" in the passage quoted above. Propositions that do not lie on the same path are probabilistically incommensurable (example: V and Z). A proposition can lie on several paths (examples: W, V) and thus be more probable than several mutually incommensurable propositions (example; W is more probable than V and Z). Only the propositions that lie on the bottom path OAI have numerical probabilities. A proposition that lacks a numerical probability can lie on a path connecting propositions with non-trivial numerical probabilities and thus be more probable or less probable than such propositions (example: V, which lies on a path leading to A). And, of course, every maximal path (i.e., a path that cannot be further extended) runs from O to I.

All these possibilities can be represented in the intersection model. Thus, it seems that the approach I have sketched is close in spirit to the Keynesian conception of probability. It takes its departure from his FA-style analysis of this concept and it builds on this analysis to provide a model that makes room for non-numerical probabilities and for various incommensurability phenomena. But it goes beyond Keynes in distinguishing between different types of incommensurability. Keynes was anxious, I think, to establish that some propositions might be probabilistically *incomparable*, because weigh-

ing different aspects of the available evidence might sometimes be impossible.[30] But he did not seem to have considered the possibility of probabilistic *parity*.

Indeed, Keynes might not have been prepared to accept the possibility that different doxastic states could be permissible. It is symptomatic that he does not distinguish between a stronger and a weaker level of normativity in his FA-style analysis of probability. He uses just one term, "rational", for the normative component. If he were confronted with this issue, he might well have argued for Uniqueness, according to which there is only one permissible doxastic state. Then it is another matter that this state could not be fully opinionated on his view. On the opposite, it would have to contain a large number of credence gaps, to account for Keynes's claims about probability. (For an interpretation of Keynes along these lines, see Weatherson 2002.)

Whether Uniqueness is a correct condition crucially depends on whether different credences in propositions that are being compared, *A* and *B*, can be permissible when

> (i) different parts of evidence point in different directions, with some parts supporting *A* and other parts supporting *B*,
>
> and
>
> (ii) there is no single way in which we should weigh the different parts of evidence against each other.

If one believes that in all such cases a credence gap with regard to *A* and *B* is mandatory, then Uniqueness is the right condition to assume. But if one allows that weight assignments to different parts of evidence might be optional to some extent and that it might therefore be permissible to give

---

[30] Here is a representative example of Keynes explicitly denying the possibility of weighing different aspects of evidence against each other: "Consider three sets of experiments, each directed towards establishing a generalisation. The first set is more numerous; in the second set the irrelevant conditions have been more carefully varied; in the third case the generalisation in view is wider in scope than in the others. Which of these generalisations is on such evidence the most probable? There is, surely, no answer […] We cannot always weigh the analogy against the induction, or the scope of the generalisation against the bulk of the evidence in support of it." (Keynes 1921, p. 31.)

more credence to *A* and likewise permissible to give more credence to *B*, then one needs a model without a unique doxastic state.[31]

Rosen (2001, p. 71) gives expression to such doxastic permissivism:

> It should be obvious that reasonable people can disagree, even when confronted with a single body of evidence. When a jury or a court is divided in a difficult case, the mere fact of disagreement does not mean that someone is being unreasonable. Paleontologists disagree about what killed the dinosaurs. And while it is possible that most of the parties to this dispute are irrational, this need not be the case. To the contrary, it would appear to be a fact of epistemic life that a careful review of the evidence does not guarantee consensus, even among thoughtful and otherwise rational investigators.

If, when confronted with the same evidence, reasonable people can disagree, then their opposing judgments, *A* and *B*, are equally permissible given the evidence in question. But then it follows that it is permissible, given this evidence, to give *A* more credence than *B* and likewise permissible to give *B* more credence than *A*. In other words, the possibility of reasonable disagreement among epistemic peers with access to the same evidence implies the falsity of Uniqueness.

Whether Uniqueness should be accepted or rejected is hotly debated at present. My own sympathies lie with the critics

---

[31] There is also another kind of cases that present difficulties for Uniqueness. Those are cases in which the agent's beliefs, whatever she decided to believe, would be self-supporting, i.e., would themselves provide evidence for their correctness. It is like a doctor's advice to a patient: "To recover you need to believe that you will; if you despair, you chances of recovery will be much lower." Whatever the patient is going to believe about her recovery will provide a support for that belief. (Cf. Drake 2017 and Reisner 2018.) But cases of this sort are very unusual, of course.
Note, by the way, that the fact that a belief in one's recovery would be self-supporting is an interesting example of a pragmatic reason to believe – an example of a a good reason to believe but still a reason of the 'wrong kind'. It is not an evidentiary reason. That this belief would be self-supporting doesn't make its content more probable; even though the belief itself would make it so. That I actually believe in my recovery *is* an evidentiary reason for what I believe.

of this constraint.[32] But I recognize I may be wrong. For values in general, permissivism seems to be very plausible. There is no need to maintain a preferential gap just because one recognizes that there is no mandatory preferential stance in a given case. But permissivism about credences is not as obviously plausible. It wouldn't be wholly surprising if probability differed from other values in this important respect. Thus, even though probability is a value, it might turn out to be a value of a very special kind – one that leaves no room for parity. The jury is still out on doxastic permissivism.

University of Lund

Acknowledgements

This paper was prepared for a meeting of Institut International de Philosophie in Helsinki in August 2019. I wish to thank the organizers of this meeting, Pascal Engel and Sami Pihlström, and the members of the Institut who were present. Prior to the meeting of the Institut, in June 2019, it was presented at a workshop on cognitive value organized in Stockholm, at the Royal Academy of Letters, by Graham Oddie and Kevin Mulligan. Then, in November the same year I presented it at an MSPT-seminar at the Research School of Social Sciences in Canberra and in December at a workshop on incommensurability and indeterminacy, at the Institute of Futures' Studies in Stockholm. In February 2020, it was discussed at a seminar in practical philosophy at Lund University. I am greatly indebted to the participants in these events and to anonymous referees for very helpful comments. Special thanks are due to Toni Rønnow-Rasmussen, with whom I have had many discussions on some of the themes of this paper.

[32] For recent criticisms of Uniqueness, see, for example, Kelly (2013), Schoenfield (2014), Meacham (2014) and Mayo-Wilson & Wheeler (2016). For defenses of Uniqueness, see White (2005, 2013), Horowitz (2014) and Hedden (2015). These are just some examples; for a helpful overview of this debate, see Kopec & Titelbaum (2016).

References

Anscombe, G. E. M. (1957), *Intention*, Basil Blackwell, Oxford.

Andersson, H. and Green Werkmäster, J. (2020), "How Valuable Is It?", published on-line *Journal of Value Inquiry*.

Boole, G. (2009 [1854]), *An investigation of the Laws of Thought, on which are founded the mathematical theories of Logic and probabilities*, Cambridge University Press, Cambridge.

Brentano, F. (1969 [1889]), *The Origin of Our Knowledge of Right and Wrong* [*Vom Ursprung sittlicher Erkenntnis*], transl. by R. M. Chisholm and E. H. Schneewind, Routledge & Kegan Paul, London and Henley.

Carabelli, A. (1998), "Keynes on Probability, Ucertainty and Tragic Choices", *Cahiers d'économie politique / Papers in Political Economy* 30/31, pp. 187-226.

Carnap, R. (1962), "The Aim of Inductive Logic", in E. Nagel, P. Suppes and A. Tarski (eds.), *Logic, Methodology and Philosophy of Science: Proceedings of the 1960 International Congress*, Stanford University Press, Stanford, pp. 303-318.

Carnap, R. (1971), "Inductive Logic and Rational Decisions", in R. Jeffrey and R. Carnap (eds.), *Studies in Inductive Logic and Probability*, vol. 1, University of California Press, Berkeley and Los Angeles, pp. 5-31.

Chang, R. (1997), "Introduction", in R. Chang (ed.), *Incommensurability, Incomparability and Practical Reason*, Harvard University Press, Cambridge, MA, pp. 1-34.

Chang, R. (2002a), *Making Comparisons Count*, Routledge, London.

Chang, R. (2002b), "The Possibility of Parity", *Ethics* 112, pp. 659-688.

Chang, R. (2005), "Parity, Interval Value, and Choice", *Ethics* 115, pp. 331-350.

Chang, R. (2019), "Defusing Continua Arguments, or, How to Avoid the Repugnant Conclusion", draft.

Cournot, A. A. (1843), *Exposition de la théorie des chances et des probabilités*, L. Hachette, Paris.

Drake J, (2017), "Doxastic Permissiveness and the Promise of Truth", *Synthese* 194, pp. 4897– 4912.

De Morgan, A. (1847), *Formal Logic: or, The Calculus of Inference, Necessary and Probable,* Taylor and Walton, London.

Donkin, W. F. (1851), "On Certain Questions Relating to the Theory of Probabilities", *Philosophical Magazine*, Series 4, vol. I, No. 5, pp. 353-368.

Eriksson, L., and W. Rabinowicz (2012), "The Interference Problem for the Betting Interpretation of Degrees of Belief", *Synthese* 190, pp. 809–30.

Fine, K. (1975), "Vagueness, Truth and Logic", *Synthese* 30, pp. 265-300.

Galavotti, M. C., (2011), "The Modern Epistemic Interpretations of Probability: Logicism and Subjectivism", in D. M. Gabbay, S. Hartmann and J. Woods (eds.), *Handbook of the History of Logic*, vol. 10: *Inductive Logic*, North Holland, Amsterdam, pp. 153-204.

Griffin, J. P. (1986), *Well-being: Its Meaning, Measurement and Moral Importance*, Clarendon Press, Oxford.

Hedden, B. (2015), "Time-slice Rationality", *Mind* 124, pp. 449-491.

Horowitz, S. (2014), "Immoderately Rational", *Philosophical Studies* 167, pp. 41-56.

Humberstone, I. L. (1992), "Direction of Fit", *Mind* 101, pp. 59-84.

Jeffreys, H. (1931), *Scientific Inference*, Cambridge University Press, Cambridge.

Jeffreys, H. (1939), *Theory of Probability*, Clarendon Press, Oxford.

Jevons, W. S. (1873), *The Principles of Science*, Macmillan, London.

Kelly, T. (2013), "Evidence Can Be Permissive", in M. Steup, J. Turri and E. Sosa, *Contemporary Debates in Epistemology*, 2nd edition, Wiley-Blackwell, pp. 298-311.

Kopec, M. and Titelbaum, M. (2016), "The Uniqueness Thesis", *Philosophy Compass* 11, pp. 189-200.

Lewis, D. K. (1980), "A Subjectivist's Guide to Objective Chance", in R. C. Jeffrey (ed.), *Studies in Inductive Logic and Probability*, Vol. II, University of California Press, Berkeley, pp. 263-293.

Lewis, C. T. & Short, C. (1879), *A Latin Dictionary*, Oxford. Clarendon Press. Keynes, J. M. (1921), *A Treatise on Probability*, Macmillan, London.

Mayo-Wilson, C. and Wheeler, G., (2016), "Scoring Imprecise Credences: A Mildly Immodest Proposal", *Philosophy and Phenomenological Research* 93, pp. 55-78.

Meacham, C. J. G. (2014), "Impermissive Bayesianism", *Erkenntnis* 79, 1185–1217. Moore, G. E. (1993 [1903]), *Principia Ethica,* Cambridge University Press, Cambridge.

Moore, G. E. (1903), "Book Review: *The Origin of The Knowledge of Right and Wrong*. Franz Brentano", *The Journal of Ethics* 14, pp. 115-123.

Poisson, S.-D. (2013 [1837]) *Researches into the Probabilities of Judgements in Criminal and Civil Cases* [*Recherches sur la probabilité des jugements en matière criminelle et en matière civile*], Paris, transl.by O. Sheynin, NG-Verlag, Berlin.

Rabinowicz, W. (2008), "Value Relations", *Theoria* 74, pp. 18-49.

Rabinowicz, W. (2012), "Value Relations Revisited", *Economics and Philosophy* 28: 133-64. Rabinowicz, W. (2017), "From Values to Probabilities", *Synthese* 194, pp. 3901-3929.

Rabinowicz, W. & Rønnow-Rasmussen, T. (2004), "The Strike of the Demon: On Fitting Pro- attitudes and Value", *Ethics* 114, pp. 391 - 423.

Raz, J. (1986), "Incommensurability", in J. Raz, *The Morality of Freedom*, Oxford University Press, Oxford, pp. 321-368.

Reisner, A., (2018), "Pragmatic Reasons for Belief", in *The Oxford Handbook of Reasons and Normativity* ed, ed.by D. Star, Oxford University Press, Oxford.

Rosen, G. (2001), "Nominalism, Naturalism, Epistemic Relativism", *Philosophical Perspectives* 15, pp. 69-91.

Schoenfield, M. (2014), "Permission to Believe: Why Permissivism Is True and What It Tells Us About Irrelevant Influences on Belief", *Nous* 48, pp. 193-218.

Skorupski, J. (2010). *The Domain of Reasons*, Oxford University Press, Oxford.

Temkin, L. (2012), *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*, Oxford University Press, Oxford.

Weatherson, B. (2002), "Keynes, Uncertainty and Interest Rates", *Cambridge Journal of Economics* 26, pp. 47-62.

Weatherson, B. (2007), "The Bayesian and the Dogmatist", *Proceedings of the Aristotelian Society* 107, pp. 169–185.

White, R. (2005), "Epistemic Permissivism", *Philosophical Perspectives* 19, pp. 445-459.

White, R. (2013), "Evidence Cannot Be Permissive", in M. Steup, J. Turri and E. Sosa, Contemporary Debates in Epistemology, 2nd edition, Wiley-Blackwell, pp. 312-323.

Zabell, S. L. (2011), "The Subjective and the Objective", in P. S. Bandyopadhyay and M. R. Forster (eds.), *Philosophy of Statistics*, North Holland, Amsterdam, pp. 1149-1174.

Non-Modal Normativity and Norms of Belief

TIMOTHY WILLIAMSON

1. Introduction

We are used to expressing and conceiving norms in *deontic modal* terms. They concern what one *may* or *must* (or *should*, or *ought* or *has* to) do or not do, be or not be. What one may do is *permissible*; what one must do is *obligatory* (not permissible not to do); what one must not do is *impermissible* or *forbidden* (not permissible to do). The semantic study of modal operators in natural language constitutes a mature research programme, whose explanatory power has been applied to understanding how such deontic modal terms work, and what logical principles they obey (Kratzer 1977, 1981, 2012; Portner 2009; Charlow and Chrisman 2016).

But, this paper argues, there is a catch. Deontic modals are less perspicuous than they seem for articulating norms. The reason is that, as modal operators, they have a dimension of complexity *extrinsic* to the norms. In practice, this extra dimension has sometimes distracted theorists of norms, and sent them off on wild goose chases. We do better expressing and conceiving the norms in simpler, more perspicuous, non-modal terms. Although we can still recapture the corresponding deontic modal operators, that is a sideshow. The second half of the paper applies this framework to clarify some confusions which have sent the debate on norms of belief off on some unrewarding detours.

2. Compliance and permissibility

Here is a simple norm, a rule, expressed sub-sententially:

Rule R No dancing in the library.

To keep things simple, we consider normativity just with respect to rule R, ignoring all other rules or norms. For example, although making a rule such as R inevitably has moral consequences, we bracket them. R is more similar to a legal rule than to a moral one. In realistic cases, norms of several kinds and different priorities interact in various complex ways. Such interactions constitute an important topic in their own right, but for present purposes it is best to screen out such effects, in order to isolate an issue which does not depend on them.

A rule such as R induces a binary distinction between the *permissible* and the *impermissible*. Most obviously, it makes dancing in the library impermissible. For present purposes, we can ignore more finely graded normative distinctions between better and worse.

2.1 Permissibility in context

Consider Alexandria, who is working in the library, not dancing. That is permissible. But is it permissible for Alexandria to dance? The natural answer is 'No—she's in the library'.

Here is another question, also asked while Alexandria is working in the library, not dancing: is it permissible for Alexandria to leave the library and dance? The natural answer is 'Yes—the rule says nothing about what you can do outside the library'.

Now we have a puzzle, for we have rejected (1) and accepted (2):

(1) It is permissible for Alexandria to dance.

(2) It is permissible for Alexandria to leave the library and dance.

In other words, though Alexandria may leave the library and dance, she must not dance. But, given standard semantic treatments of the two sentences, (2) logically entails (1). More generally, standard semantic accounts validate these forms of argument:

P(&E) From 'It is permissible for S to φ and ψ', derive 'It is permissible for S to φ' and 'It is permissible for S to ψ'.

We can put the reason like this, ignoring various subtleties in the semantics which do not matter here. The premise means that S φs and ψs in some relevant permitted situation *s*, where a situation is permitted just in case it contains no violation of R. Suppose that the premise is true. Then S φs in *s* and S ψs in *s*. So S φs in some relevant permitted situation and S ψs in some relevant permitted situation. But those are what the conclusions mean. Thus both conclusions are true. Therefore P(&E) is guaranteed truth-preserving.

The situation *s* need not be actual. It may be merely possible, in a suitably broad sense of 'possible'.

The treatment of 'permissible' as an existential quantifier is natural, and standard in deontic logic. After all, if it is *im*permissible for S to φ, then no relevant situation in which S φs is permitted. Conversely, if it is permissible for S to φ, then it is permissible for S to φ in some specific way, so the situation of S φing in that way is permitted. In terms of modal logic, P(&E) corresponds to the inference from $\Diamond(\alpha \,\&\, \beta)$ to $\Diamond\alpha$ or $\Diamond\beta$, which is valid in any normal modal logic, and indeed in any monotonic modal logic (Chellas 1980).

Technically, one can devise non-monotonic deontic logics in which $\Diamond(\alpha \,\&\, \beta)$ does not in general entail either $\Diamond\alpha$ or $\Diamond\beta$, and even provide them with a possible-worlds model theory of sorts in the framework of *neighbourhood semantics* (again, Chellas 1980). The reason for not going that way is explained below.

The motivation for P(&E) was put in terms of possible situations, rather than possible worlds, because it takes much less for a possible situation to avoid violations of the rule than it does for a possible world to avoid them: worlds are global, situations are local. For example, imagine that in all the relevant worlds Jane obsessively dances in the library, so R is always violated. That should not affect the permissibility of *Alexandria's* dancing. We can achieve this effect by taking account of possible small situations containing Alexandria's actions but not Jane's, so they need not contain violations of R. Possible worlds semantics can achieve the same effect, but less directly. Since most semantic accounts of modals use

worlds rather than situations, we usually do the same in this paper, for ease of comparison, and achieve the 'localizing' effect of situations by explicit relativization to specific features of the situation, such as the agent.

Someone might object that the argument from (2) to (1) is not a genuine instance of P(&E), because it equivocates about the time of the dancing: (1) concerns what it is permissible for Alexandria to do *now*, while (2) concerns what it is permissible for her to do *later*. But time is not really the key to the problem. Both (1) and (2) are naturally understood as about the near future, not the present instant. When we ask 'Is it permissible for Alexandria to dance?', our interest is not always in whether it is permissible for her to be *already* dancing. For example, it may be common ground in the conversation that some dance music will start in two minutes, at noon, loud enough to be heard both inside and outside the library. Stopping work and leaving the library will take Alexandria less than two minutes. Thus the time of the dancing in both (1) and (2) may be naturally understood as starting from noon.

Nevertheless, such examples pose no threat to the validity of P(&E). For that means only that if the premise is true as uttered in a given context, the conclusions are true as uttered in the *same* context. When we originally considered (1), we held Alexandria's location in the library fixed—it was not being questioned. When we went on to consider (2), we no longer held her location in the library fixed, for it was explicitly at issue. Once we have accepted (2), it is natural for us to change our answer to the original question 'Is it permissible for Alexandria to dance?', saying 'Yes', though perhaps adding by way of explanation 'but first she must leave the library'. This fits the semantics of 'permissible' sketched above. Situations in which she left the library were implicitly excluded as irrelevant in the original context, but not in the later context. In any context in which (2) is true, (1) is also true. Standard semantic accounts of modals allow for just such context-sensitivity: the *relevant* possible worlds or situations are those in the contextually determined *modal base*, to use Kratzer's terminology.

The reversal of attitude to (1) when (2) comes into play is a typical contextual effect. It is a special case of the more gen-

eral conversational phenomenon by which it is easier to expand the domain of contextually relevant items than to contract it again. In this case, the items are possible situations or worlds. Once one has accepted (2), to continue plainly rejecting (1) suggests a conversational tin ear. A similar effect has been observed when extra conjuncts are added to the antecedent of a counterfactual conditional; the analogous pairs to (2) and (1) are known as *reverse Sobel sequences*, and the effect is evidence of context-sensitivity in the semantics of such conditionals (von Fintel 2001; Gillies 2007; for a response to Moss 2012's attempt to explain away the effect in epistemic terms see Williamson 2020, pp. 222-228). Hence (1)'s loss of acceptability on the introduction of (2) is explicable on pragmatic grounds, as a special case of an independently attested effect, with no threat to the structurally attractive principle P(&E). By contrast, using neighbourhood semantics to invalidate P(&E) leaves the dynamic effect unexplained. Thus the evidence strongly favours a contextual account of the apparent gap between (1) and (2).

Contextual variation has been much debated in the recent literature on 'ought' and other deontic modals (Worsnip (2019) provides a helpful survey). However, the focus has not been on variation in the modal base. Instead, it has concerned what *kind* of normativity is intended—moral, legal, prudential, …—and whether the relevant *standards* vary with the values of participants in the conversation. In Kratzer's terminology, the variation at issue is in the *ordering source* for 'ought'. Variation in the modal base has attracted much less attention.

Context-sensitivity in the modal base is commonplace in natural language. But it has non-trivial methodological consequences for inquiry into norms. Consider a theorist who seeks a general necessary and sufficient condition for some sort of permissibility, perhaps for the toy example where only rule R is in play. If the candidate condition is formulated in contextually invariant terms, it will fall to counterexamples in at least some contexts, since 'permissible' is contextually variable: any expression cross-contextually equivalent to a contextually invariant expression is itself contextually invariant. On the other hand, if the candidate condition is formulated in contextually variable terms, then which condition is being

discussed varies independently of the norm(s) under discussion. For the content of the theoretical formulation varies with the second-order context in which the *theorists* are discussing it, rather than with the first-order context in which those committed to the norm apply it. Of course, one can switch to a meta-linguistic formulation, which mentions the deontic modals rather than using them, and explicitly relativizes to a context of utterance. But that will not happen unless the need for it is recognized. Moreover, such metalinguistic formulations are often a complicating distraction from the meta-normative rather than meta-linguistic question originally at issue.

For example, rule R just says 'No dancing in the library': whether a given possible situation or world contains a violation of R does not vary with a contextually determined modal base. Although R does contain contextually variable elements, such as the reference of 'the library', they depend on the context where R is propounded, not on the context where 'permissibility' with respect to R is discussed—the latter is a different level of context-sensitivity.

2.2 Compliance conditions

We can display the non-modal nature of rule R by articulating a compliance condition for R. To avoid the complexities of situation semantics, while achieving a similar localizing effect, we work instead with triples of a subject S, a time *t*, and a world *w*. The condition is simple:

COMPLIANCE S complies with R at *t* in *w* if and only if S is not dancing in the library at *t* in *w*.

We abbreviate 'S complies with R at *t* in *w*' as 'Complies$_{t,w}$(S)'. To avoid clutter, the norm parameter 'R' is left implicit; it can be understood below as supplied by context.

Evidently, the compliance condition itself does not depend on a modal base. Of course, COMPLIANCE is intended to hold counterfactually, as well as actually, with the content of rule R being kept fixed, but the same holds for any other claim of necessity and sufficiency: it has nothing specific to do with the condition for which necessity and sufficiency are being claimed. Neither the left-hand side nor the right-hand

side of COMPLIANCE expresses a modal condition in any distinctive sense. There is nothing especially modal about not dancing in the library. More precisely, the condition for compliance at a time *t* in a world *w* with respect to R, 'S is not dancing in the library at *t* in *w*', involves no *quantification* over worlds, whereas the conditions for permissibility and impermissibility at *t* in *w* with respect to R *do* involve quantification over worlds.

We can understand *violating* a rule as simply not complying with it. Thus:

VIOLATION	S violates R at *t* in *w* if and only if S is dancing in the library at *t* in *w*.

Obviously, violation is just as non-modal as compliance.

Naturally, some rules have a specifically modal content. An example is RM:

RM	Make as little noise as possible!

The condition for complying with RM is to make as little noise as possible; the condition for violating RM is to not make as little noise as possible. Those are both modal conditions, but that just concerns the specific content of RM, not its general status as a rule.

Of course, the mere statement that Alexandria complied with rule R is normatively neutral, in the sense that making it does not imply that R has any sort of normative authority. The same goes for the statement that she violated R. Once one accepts R's authority, one needs some way to express it, and the normative positions it puts agents in. The salient worry is that we need deontic modals for that purpose, for example to say 'She ought to comply with the rule', so that falling back on compliance and violation avoids the complexities of deontic modals only at the cost of giving up normativity altogether.

In response, an initial point is that the normative nature of deontic modals is independent of their contextual variation with the modal base; the variation is neither necessary nor sufficient for the normativity. It is not sufficient, because non-deontic modals such as the 'can' of ability also vary contextually with the modal base. It is not necessary, because holding the modal base fixed—for example, to contain all metaphysi-

cally possible worlds—would not somehow neutralize the normative force of deontic modals. In principle, therefore, a term could do the normative work without the contextual variation. Indeed, we can judge someone's past action or omission 'right' or 'wrong', 'appropriate' or 'inappropriate', 'correct' or 'incorrect', 'OK' or 'not OK', in virtue of its having complied with or violated a norm, without needing a modal base to make sense of the judgment. For instance, after Alexandria danced in the library, the librarian may judge her action 'wrong', 'inappropriate', 'incorrect', or 'not OK'. Thus we can put a normative spin on compliance and violation without recourse to deontic modals.

2.3 Recovering permissibility from compliance

Given a normative standard of compliance, we can then combine it with a contextually specified modal base to recapture the corresponding deontic modality. Since compliance and violation have been characterized in effect over triples <S, *t*, *w*> of a subject, a time, and a world, we treat contextual relevance as a constraint on such triples. But instead of excluding some triples as absolutely irrelevant in the context, outside the domain of quantification (the analogue of the modal base), we treat contextual relevance as a contextually determined binary relation between triples. For example, consider (3), uttered in a context where it expresses a truth:

(3) It is not permissible for Alexandria to dance, but if she were outside the library it would be permissible for her to dance.

The first conjunct treats possible situations where Alexandria is outside the library as contextually irrelevant to permissibility in the actual situation, while the second conjunct treats them as contextually relevant to permissibility in a counterfactual situation where she is already outside the library. Similarly, consider (4), uttered in a context where it expresses a truth:

(4) It is not permissible for Alexandria to dance, but before she entered the library it was permissible for her to dance.

The first conjunct treats past situations when Alexandria was outside the library as contextually irrelevant to present per-

missibility, while the second conjunct treats them as contextually relevant to past permissibility before she entered the library. Since the relevance comes from sameness in *Alexandria's* location, and in the conversation we may be contrasting what is permissible for her with what is permissible for other subjects, we also need the subject parameter S in the triples. However, we treat that parameter differently from the time and world parameters *t* and *w*, holding the former fixed in the semantic clauses, when it is explicitly specified in the sentence, while letting the implicit parameters vary. A counterpart theorist such as David Lewis would let the explicit parameter vary too.

Beyond these bare structural constraints, we leave contextual relevance unanalyzed, since trying to give necessary and sufficient conditions for such messy pragmatic relations is a mug's game, and in any case not to the present purpose.

Here is the result, with '$\varphi_{t^*,w^*}(S)$' abbreviating 'S φs at *t** in *w**', and 'contextually relevant' occurring unanalyzed on the right-hand side:

PERMISSIBILITY It is permissible with respect to R at *t* in *w* for S to φ if and only if for some triple <S, *t**, *w**> contextually relevant to <S, *t*, *w*> such that $\text{Complies}_{t^*,w^*}(S)$, $\varphi_{t^*,w^*}(S)$.

Here *t** and *w** are a time and world at which the putatively permitted action is performed, whereas *t* and *w* are a time and world at which it is putatively permissible: as already noted, when we ask whether something is permissible, we often have in mind doing it in the near future rather than already doing it, and many things are permissible without actually being done. We abbreviate the left-hand side as '$\text{Permissible}_{t,w}(\varphi(S))$'.

In any context where relevance is a reflexive relation, PERMISSIBILITY makes compliance and performance together imply permissibility. More precisely, $\text{Permissible}_{t,w}(\varphi(S))$ whenever $\text{Complies}_{t,w}(S)$ and $\varphi_{t,w}(S)$ because <S, *t*, *w*> is contextually relevant to itself.

It is also easy to check that PERMISSIBILITY validates the principle P(&E), for virtually the same reason as before.

We can define obligation as the dual of permissibility in the usual way:

OBLIGATORINESS It is obligatory with respect to R at t in w for S to φ if and only if for every triple <S, t^*, w^*> contextually relevant to <S, t, w> such that $\text{Complies}_{t^*,w^*}(S)$, $\varphi_{t^*,w^*}(S)$.

We abbreviate the left-hand side of OBLIGATORINESS as 'Obligatory$_{t,w}$(φ(S))'. Thus, with ¬ for negation, Obligatory$_{t,w}$(φ(S)) is equivalent to ¬Permissible$_{t,w}$(¬φ(S)) and Permissible$_{t,w}$(φ(S)) to ¬Obligatory$_{t,w}$(¬φ(S)).

As expected, in any context where relevance is reflexive, compliance and obligation together imply performance. More precisely, $\varphi_{t,w}(S)$ whenever $\text{Complies}_{t,w}(S)$ and Obligatory$_{t,w}$(φ(S)).

In the usual way, we can make the time-world parameters implicit in a tensed modal language where formulas are evaluated at time-world pairs, and contextual relevance is treated like an accessibility relation between such pairs. However, in line with the earlier discussion, we keep an explicit parameter S for the subject, subscripted to the deontic modal operators, and prefix it to the world-time pairs to use the original relation of contextual relevance between subject-time-world triples. To emphasize the modal analogies, we use $\Diamond_S$ for 'it is permissible for S' and $\Box_S$ for 'it is obligatory for S', formalizing 'It is permissible for S to φ' as $\Diamond_S$ φ(S) and 'It is obligatory for S to φ' as $\Box_S$ φ(S). The semantic clauses for the operators are then:

$\Diamond_S$ α is true at <t, w> if and only if for some t^* and w^* such that <S, t^*, w^*> is contextually relevant to <S, t, w> and $\text{Complies}_{t^*,w^*}(S)$: α is true at <$t^*$, w^*>.

$\Box_S$ α is true at <t, w> if and only if for every t^* and w^* such that <S, t^*, w^*> is contextually relevant to <S, t, w> and $\text{Complies}_{t^*,w^*}(S)$: α is true at <$t^*$, w^*>.

As one can easily check, once 'Complies' was introduced, the rest of the development was rather general. Although PERMISSIBILITY and OBLIGATORINESS include the phrase 'with respect to R', nothing in their structure depends on the very specific content of rule R. Many other rules or pluralities of rules would do instead, with corresponding understandings of 'Complies'.

A further complexity is that whether a rule such as R is in force is itself contingent and temporary: R had to be introduced. Universal morality does not forbid dancing in the library. Yet OBLIGATORINESS makes compliance obligatory (with respect to R) at all times in all worlds: if α is true at exactly the pairs <*t*, *w*> for which $\text{Complies}_{t,w}(S)$, then trivially $\Box_S$ α is true at any time-world pair whatsoever. In effect, the semantics has treated rule R as absolutely given. To take account of the contingent and temporary nature of such norms, we can consider a higher-order norm: to comply with any norm relevantly similar to R and currently (temporarily, contingently) in force. That is itself a higher-order compliance condition, and a similar semantic development can be carried through at this higher level. However, to simplify the exposition, that extra dimension of complexity will be ignored in what follows. It can easily be reintroduced.

2.4 How local is compliance?

Unsurprisingly, one can construct tricky cases for the semantic framework just sketched of rules whose violations are hard to *localize*. Here is an example on the temporal dimension. Imagine a lax train company with this rule:

RT Any journey made without a ticket must subsequently be paid for.

There is no deadline. Even after the traveller's death, someone else can pay for the journey. Although RT can be violated, if a journey is never paid for, it is hard to say *when* the violation occurs—at every time after the journey has started in such a world? At worst, one might have to fall back on a coarser-grained version of the semantics which eliminates the temporal parameter from compliance.

Similarly, violations are sometimes hard to localize on the dimension of the subject S. An office may have this rule:

RS At least one member of staff must be present on the mezzanine level at any time.

When no member of staff is present on the mezzanine level, RS is violated, but there may be no particular member of staff who is violating it. Perhaps one should just say that the

members of staff collectively are violating RS, and allow plural subjects in the semantics.

How far the distinction between compliance and violation can be localized depends on the content of the rule. That complicates the picture. Some flexibility is needed in applying the theoretical template; it cannot always be done mechanically. But that is not to say that it sometimes cannot be applied at all. The non-modal distinction between compliance and violation still takes priority over the modal distinction between permissibility and impermissibility.

2.5 The normativity of compliance

Another worry about RS and RT is that they are both stated using a modal operator, 'must', read deontically. For this point, the difference between 'must' and 'obligatory' does not matter; restating RT and RS in terms of 'obligatory' makes little difference. If the rules which lay down the distinction between compliance and violation to begin with are formulated in terms of deontic modals, how can the distinction take priority over those very deontic modals?

A short response is that we need the distinction between compliance and violation to make sense of the deontic modals. This is not simply an appeal to the definitions just proposed. Start with the unanalyzed deontic modals themselves, and put yourself in an extreme context focused just on living in the present and the actual, with no interest in cross-time or cross-world comparisons. In this case, contextual relevance between triples is simply identity. By the standards of this context, if what one is doing in a situation counts as permissible for one, in that situation one is complying with the relevant norms; if what one is doing in the situation does not count as permissible for one, one is violating some of those norms. Thus the distinction between compliance and violation can be recovered from deontic modals in a suitable context.

In less favourable contexts, the distinction may not be recoverable, because some situations are ignored as entirely irrelevant in the application of deontic modals, so nothing can be gathered from the application of the deontic modals as

to whether those situations comply with the norms or violate them.

Once we have the distinction between compliance and violation, we can put a suitable normative spin on it, as noted above. For example, we can use 'OK' for the desired normative spin, without making it a deontic modal. Then we can informally paraphrase RT and RS:

RTa	Journeys made without a ticket and never subsequently paid for are not OK.
RSa	Situations when no member of staff is present on the mezzanine level are not OK.

Although RTa and RSa lack the pomposity characteristic of regulations, they do the job.

In any case, the role of deontic modals in stating rules is not to introduce context-dependence; it is to make explicit the rules' intended normative force. The context-dependence simply comes along for the ride. Indeed, the semantics above for the obligation operator $\Box_S$ makes '$\Box_S$ (S complies)' true under any contextual restriction for the standard of compliance at issue; the same applies to statements of rules in deontic modal form. The contextual variability of deontic modals can be compared to the ever-present possibility of implicit contextual restrictions on the domains of quantifiers in natural language. That does not stop us from using those quantifiers as a convenient way of stating universal generalizations for which only the explicit restrictions are intended, without exploiting their potential for being implicitly contextually restricted; if absolutely every *F* is *G*, then *a fortiori* every *F* in a contextually restricted domain is *G*. We may likewise use deontic modals as a convenient way to make norms explicit, without exploiting their similar potential for implicit contextual restriction.

3. Application to norms of belief

The considerations in section 2 alert us to problematic assumptions in the current debate on norms of belief (e.g. Chan 2013, Matheson and Vitz 2014). Norms of belief in turn cast light on problems in localizing violations of norms.

Belief is subject to norms: one can believe something rightly or wrongly, correctly or incorrectly, reasonably or unreasonably, rationally or irrationally, with or without justification, and so on. Some of those distinctions focus more on the content of the belief, others more on how it was formed or is maintained. Some philosophers speak of doxastic or epistemic virtues or vices, but there is no need to moralize: basic normative distinctions amongst beliefs apply even to the beliefs of non-human animals and very young children, to which a moralizing approach is quite inappropriate. A norm for belief is more fruitfully treated as distinguishing between *defective* and *non-defective* beliefs, perhaps in a broadly functional sense (see Williamson forthcoming for discussion). Identifying a basic norm of belief thus becomes a way of understanding the *nature* of belief—which supports an approach applicable to *all* beliefs, not just to those of conscious, responsible agents.

3.1 Some salient norms of belief

One salient candidate norm of belief is the *truth norm*. We can put it in the imperative:

TNB Believe *p* only if *p* is true!

If we treat TNB as the basic norm of belief, we can extract a corresponding standard of non-defectiveness from it:

TND If S believes *p*, S's belief in *p* is non-defective if and only if *p* is true.

A more demanding candidate norm of belief is the *knowledge norm*:

KNB Believe *p* only if you know *p*!

Since true belief does not entail knowledge, complying with TNB does not entail complying with KNB. But since knowledge does entail truth, complying with KNB does entail complying with TNB. However, treating KNB as the basic norm of belief is incompatible with treating TNB as its basic norm, and the corresponding standard of non-defectiveness is incompatible with TND (they cannot both be necessary and sufficient for non-defectiveness in the same sense):

KND — If S believes *p*, S's belief in *p* is non-defective if and only if S knows *p*.

Another less demanding candidate norm of belief is the *evidence norm*:

ENB — Believe *p* only if you have good evidence for *p*!

'Good evidence' is so understood here that one can count as having good evidence for a falsehood. Thus complying with ENB does not entail complying with TNB, and so *a fortiori* does not entail complying with KNB. We also understand 'good evidence' so that one counts as having good evidence for *p* whenever one knows *p*. Thus complying with KNB entails complying with ENB. However, treating ENB as the basic norm of belief is incompatible with treating either TNB or KNB as its basic norm, and the corresponding standard of non-defectiveness is incompatible both with TND and with KND:

END — If S believes *p*, S's belief in *p* is non-defective if and only if S has good evidence for *p*.

Of course, one can construct many other candidate norms of belief, even candidate basic norms of belief. This paper is not concerned with deciding between these candidates, but with the structure of the issues.

3.2 The permissibility of belief and Moore paradoxes

Some participants in the debate on norms of belief insist that a proper account of them must answer the question 'When is it permissible to believe *p*?' They want a necessary and sufficient condition for the permissibility of belief. Clearly, not even a necessary and sufficient condition for non-defectiveness such as TND, KND, or END specifies a necessary and sufficient condition for permissibility, since they apply only when S *does* believe. Even when S lacks belief in *p*, it may still be *permissible* for S to believe *p*, but a condition for non-defectiveness does not say when.

Of course, one can propose conditions for permissibility mechanically corresponding to conditions for non-defectiveness:

TNP	It is permissible for S to believe *p* if and only if *p* is true.
KNP	It is permissible for S to believe *p* if and only if S knows *p*.
ENP	It is permissible for S to believe *p* if and only if S has good evidence for *p*.

These are *narrow-scope* norms: the scope of 'permissible' extends only as far as 'believe *p*'. But TNP, KNP, and ENP are no mere reformulations of the previous principles: they introduce special problems of their own.

A much-discussed example in the debate applies the truth norm TNP to Moore paradoxes (Bykvist and Hattiangadi 2007). Imagine that it is raining, but for some reason I fail to believe that it is raining. Let *p* be the conjunction that it is raining and I fail to believe that it is raining. Thus *p* is true. Hence, by the right-to-left direction of TNP, it is permissible for me to believe *p*, in other words, permissible for me to believe that it is raining and I fail to believe that it is raining. But whenever I believe the conjunction, I believe its first conjunct, so the second conjunct is false, so the whole conjunction is false. Thus it is permissible for me to believe *p* even though, whenever I believe *p*, *p* is false. That is obviously quite at odds with the spirit of the truth norm of belief. In response, philosophers have proposed inserting various epicycles into TNP.

The difficulty comes from the modal nature of permissibility: the possible situation where it is permissible to believe *p* differs from the possible situation where *p* is believed. Recall Alexandria, sitting in the library. In the spirit of the truth norm of belief, is it permissible for her to believe that she is dancing? The natural answer is 'No', for the belief would be false. But, still in the spirit of the truth norm (and no longer worrying about rule R), is it permissible for Alexandria to start dancing and believe that she is dancing? The natural answer is 'Yes', for the belief would be true. We again have a puzzle, for we have rejected (5) and accepted (6):

(5) It is permissible for Alexandria to believe that she is dancing.

(6) It is permissible for Alexandria to start dancing and believe that she is dancing.

But (6) entails (5), just as (2) entails (1) in the original puzzle, by principle P(&E). For the same sort of reason as in section 2, there need be no equivocation about when she putatively believes that she is dancing: the example can be set up to make the time exactly the same for (5) and (6).

Just as before, the puzzle is an artefact of contextual variation in the modal base for the deontic modal 'permissible'. In the original context when we assessed (5), we kept Alexandria's bodily movements fixed, and treated possible situations where she is dancing as irrelevant. In the context when we assess (6), we no longer keep Alexandria's bodily movements fixed, because they are explicitly at issue. We treat possible situations where she is dancing as relevant. We revise our attitude to (5) correspondingly. To the question 'Is it permissible for Alexandria to believe that she is dancing?', we might now answer 'Yes', adding by way of explanation 'but first she must start dancing'.

Complicating TNP with epicycles to handle tricky propositions *p* is therefore a waste of time. For no such epicycles will handle the contextual variability of 'permissible', which depends not only on *p* but on many other factors too, including which other propositions happen to be salient in the context. As usual, attempts to handle pragmatic phenomena in the semantics end badly.

Naturally, the issue is not confined to the truth norm; it arises for other putative norms of belief too. KNP and ENP are just as vulnerable as TNP to contextual variability. Indeed, the same example works for any reasonable norm of belief. In the possible situations relevant for the original context, not only is it not true that Alexandria is dancing; she does not know that she is dancing, has no good evidence that she is dancing, and so on. By contrast, in some of the possible situations relevant for the later context, not only is it true that Alexandria is dancing; she knows that she is dancing, has good evidence that she is dancing, and so on. Thus the contextual variation in the truth-value of (5) will occur on just about any reasonable norm of belief.

3.3 Compliant belief

We therefore have good reason to apply the framework of section 2 to norms of belief. As usual, the distinction between compliance and violation is primary. Norms such as TNB/TND, KNB/KND, and ENB/END are clear about what constitutes a violation: believing without truth, without knowledge, and without good evidence, respectively. That is what really matters. The rest is just a matter of finding the best way to fit deontic modals on top of that.

However, we must be careful how we do it. Let us try applying PERMISSIBILITY just as it stands in section 2. That requires cashing out '$\text{Complies}_{t,w}$(S)' for a given norm of truth. What is it for a subject S to comply with such a norm at a time t in a world w? For the truth norm, the natural answer is this: for every proposition p such that S believes p at t in w, p is true. For the knowledge norm, it is: for every proposition p such that S believes p at t in w, S knows p at t in w. For the evidence norm it is: for every proposition p such that S believes p at t in w, S has good evidence for p at t in w. By plugging 'the given norm of belief' for 'R' and 'believe p' for 'φ' into PERMISSIBILITY above, we then obtain this equivalence (for the given norm of belief):

(7) It is permissible at t in w for S to believe p if and only if for some triple <S, t*, w*> contextually relevant to <S, t, w> such that $\text{Complies}_{t^*,w^*}$(S), S believes p at t* in w*.

The development can subsequently proceed as in section 2.

However, that literal application of the template faces a problem. We first consider it for the truth norm. Let p be the proposition 'Not everything I believe is true'. Suppose that I believe p. Then either p is not true, so not everything I believe is true, or p is true, so not everything I believe is true (for that is what p says). Either way, not everything I believe is true. Thus believing p guarantees that I violate the truth norm. In effect, '$\text{Complies}_{t^*,w^*}$(I)' is incompatible with 'I believe p at t* in w*'. Consequently, by (7), it is *not* permissible for me to believe p. But, whenever I believe p, since not everything I believe is true, p is true (because that is what it says). Thus my belief in p itself satisfies the truth norm, even though it guarantees that I will violate the truth norm elsewhere. There

is a defect, but not in p. The condition for permissibility in (7) fails to capture that.

The problem is not confined to the truth norm. Since knowledge entails truth, the same argument from (7) shows that my belief in p is also not permissible with respect to the knowledge norm. Yet I may even *know* p, for example because I know that I have inconsistent beliefs. Thus my belief in p itself satisfies the knowledge norm too, even though it guarantees that I will violate the knowledge norm elsewhere. Again, there is a defect, but not in p. The condition for permissibility in (7) fails to capture that too. The evidence norm may also raise similar problems.

The argument makes a contentious assumption: that the sentence 'Not everything I believe is true' expresses the same proposition as used at different times t and t^*. Arthur Prior would accept that assumption; many other philosophical logicians would not. If p is simply the proposition that not everything I believe at t is true, then the reasoning about its consequences for my beliefs at t^* does not go through. However, we can accommodate the alternative view of propositions by taking p^T to be instead the proposition that at each time in the interval T not everything I believe is true, where T includes every time t^* contextually relevant to t (with respect to some worlds). Then a variant of the reasoning still goes through. For suppose that I believe p^T at a relevant time t^* (in a world w^*). Hence either p^T is not true, in which case not everything I believe at t^* is true, or p^T is true, so at each time in T not everything I believe is true (for that is what p^T says), so not everything I believe at t^* is true (for t^* is in T). Either way, not everything I believe at t^* is true. Thus believing p^T guarantees that I violate the truth norm at t^*. In effect, 'Complies$_{t^*,w^*}$(I)' is incompatible with 'I believe p^T at t^* in w^*'. Consequently, by (7), it is *not* permissible for me to believe p^T at t. Yet p^T may well be true; indeed, it will be true if I believe it throughout the interval T. I may even *know* p^T, through my awareness of the deep-seated inconsistencies in my belief system (we all have them). We can choose a context where the interval T is short enough for such knowledge to be available, and raise the problem with respect to that context. The use of Moore's paradox as a counterexample to TNP may require similar tweaking.

A structurally similar issue arises for the norm of promise-keeping. I promise you that I will break a promise. If I keep that promise to you, I break another promise. Just by making that promise, I am guaranteed to violate the norm of promise-keeping. Nevertheless, I *can* keep the promise to you. For the norm of promising, we track the distinction between compliance and violation promise by promise, not just promiser by promiser. Similarly, for norms of believing, we track the distinction between compliance and violation belief by belief, not just believer by believer.

One way to handle the issue is by localizing compliance and violation to a proposition as well as to a subject, a time and a world. Thus we may write '$\text{Complies}_{t,w}(S, p)$' in place of '$\text{Complies}_{t,w}(S)$'. For the truth norm, the compliance condition with respect to a quadruple <S, *p*, *t*, *w*> is: if S believes *p* at *t* in *w*, *p* is true at *t* in *w* (with 'if' read materially). For the knowledge norm, it is: if S believes *p* at *t* in *w*, S knows *p* at *t* in *w*. For the evidence norm, it is: if S believes *p* at *t* in *w*, S has good evidence for *p* at *t* in *w*. And so on. We also treat the explicitly specified proposition *p* in the same way as the explicitly specified subject S with respect to the contextual relevance relation. Thus in place of (7) we have (8) (for the given norm of belief):

(8) It is permissible at *t* in *w* for S to believe *p* if and only if for some quadruple <S, *p*, *t**, *w**> contextually relevant to <S, *p*, *t*, *w*> such that $\text{Complies}_{t^*,w^*}(S, p)$, S believes *p* at *t** in *w**.

This handles the permissibility of beliefs like 'Not everything I believe is true' in the desired way.

More generally, within this framework we can define deontic modal operators $\Diamond_{S,p}$ and $\Box_{S,p}$ relativized to a subject S and proposition *p*:

$\Diamond_{S,p}$ α is true at <*t*, *w*> if and only if for some *t** and *w** such that <S, *p*, *t**, *w**> is contextually relevant to <S, *p*, *t*, *w*> and $\text{Complies}_{t^*,w^*}(S, p)$: α is true at <*t**, *w**>.

$\Box_{S,p}$ α is true at <*t*, *w*> if and only if for every *t** and *w** such that <S, *p*, *t**, *w**> is contextually relevant to <S, *p*, *t*, *w*> and $\text{Complies}_{t^*,w^*}(S, p)$: α is true at <*t**, *w**>.

In particular, we understand 'It is permissible for S to believe *p*' (or 'S may believe *p*', read deontically) as '$\Diamond_{S,p}$ (S believes *p*)'. Likewise, we understand 'It is obligatory for S to believe *p*' (or 'S ought to believe *p*') as '$\Box_{S,p}$ (S believes *p*)'.

3.4 Obligations to believe?

Some explanation is needed of the truth-condition for 'S ought to believe *p*'. As we have seen, the compliance condition for S and *p* will be of this form, for some C: either S fails to believe *p* or C(S, *p*). The truth-condition for 'S ought to believe *p*' is then that on every contextually relevant quadruple such that either S fails to believe *p* or C(S, *p*), S believes *p*. That condition fails whenever, on some relevant quadruple, S fails to believe *p*; it holds whenever on every relevant quadruple S believes *p*. It therefore simplifies to the condition that for every contextually relevant quadruple, S believes *p*. The result is independent of the specific content of the compliance condition C(S, *p*). All it depends on is the structural point that one vacuously complies by failing to believe, for the only case where a norm of the kind at issue obliges one to believe is the trivial one where failure to believe is contextually irrelevant. It is like the well-known effect of the standard deontic modal semantics that it makes all tautologies trivially obligatory: a surprising result, but the surprise is easily explicable on pragmatic grounds: there is normally no conversational point in insisting that your interlocutor *ought* to be such that 2 + 2 = 4.

The non-trivial obligations generated by such norms of belief are only to satisfy the compliance condition at issue:

TNO	S ought to (believe *p* only if *p* is true)
KNO	S ought to (believes *p* only if S knows *p*)
ENO	S ought to (believe *p* only if S has good evidence for *p*)

These are wide-scope obligations. 'S ought to' is used informally for the obligation operator $\Box_{S,p}$. Given the corresponding compliance conditions, TNO, KNO, and ENO hold under any contextual restriction.

Of course, we can also consider putative norms on which belief *is* sometimes non-trivially obligatory. These too will be derivative from a corresponding compliance condition. For example, one such norm is violated by any unbelieved truth. That norm implies a wide-scope converse obligation to TNO (under any contextual restriction):

TNO_C S ought to (believe *p* if *p* is true).

As we should expect, TNO_C does not imply the corresponding narrow-scope principle:

TNO_{CN} (S ought to believe *p*) if *p* is true.

One counterexample is a true Moore-paradoxical sentence in place of '*p*'. For since S's believing *p* falsifies *p*, S is not obliged to believe *p*; thus TNO_{CN} has a false consequent and true antecedent. But the example poses no threat to TNO_C, for S can meet that obligation simply by falsifying *p*, perhaps by believing that it is raining, contrary to the second conjunct of *p*.

A more demanding norm is violated by any *unknown* truth. Such a norm implies a wide-scope obligation connecting truth and knowledge (under any contextual restriction):

KO S ought to (know *p* if *p* is true).

Predictably, KO does not imply the corresponding narrow-scope principle:

KO_N (S ought to know *p*) if *p* is true.

A true Moore-paradoxical sentence is also a counterexample to KO but not to KO_N, for reasons similar to those for TNO_C and TNO_{CN}. A notable feature of KO is that the putative obligation does not specifically concern *belief*; it simply relates truth and knowledge.

3.5 Fine-grained belief states

Another direction of inquiry is to experiment with making the framework even more fine-grained than it already is. For we can make some sense of the idea that a given subject at a given time has *several* beliefs in the same proposition. For example, on a directly referential view of proper names, the sentences 'Hesperus is bright' and 'Phosphorus is bright' may express the same proposition. Someone who doubts the identity sentence 'Hesperus = Phosphorus' may believe a single proposition twice over, once under the mode of presentation 'Hesperus is bright' and once under the mode of presentation 'Phosphorus is bright'. We could regard him as having two beliefs, psychologically individuated, in that proposition. That would make no difference to the truth norm, since the two beliefs would have the same truth-value. But it *would* make a difference to many other putative norms of belief. For example, someone who accepts 'Hesperus = Phosphorus' on inadequate evidence may have two beliefs in the identity proposition: perhaps her belief in it under the mode of presentation 'Hesperus = Hesperus' constitutes knowledge, while her belief in it under the mode of presentation 'Hesperus = Phosphorus' does not. One belief complies with the knowledge and evidence norms while the other violates them, even though they are beliefs in the same proposition. Psychological compartmentalization, a 'divided mind', may produce similar results even without a difference in mode of presentation. We can adapt the present approach to beliefs so individuated by reinterpreting the same formal framework, by understanding the variable '*p*' as taking beliefs rather than propositions as values and tweaking the interpretation of 'S believes' accordingly. Many variations can be played on the same theme.

4. Conclusion

How far the distinction between compliance and violation can be localized depends on the specific nature of the norm at issue. For present purposes, we can leave that matter schematic, rather than seeking explicit uniformity. The moral to take away is that, when our interest is in the normative questions rather than the semantic ones, we should redirect our

focus away from the modal distinction between permissibility and impermissibility, towards the non-modal distinction between compliance and violation.

University of Oxford

Acknowledgements

This paper was presented at the 2019 Helsinki workshop on Normativity. For valuable feedback I thank audiences there and at the 2017 Whitehead Lectures at Harvard, the 2018 Paolo Casalegno Lectures at the University of Milan, a conference of the European Epistemology Network in Amsterdam, and at Oxford University, Yale University, Durham University, and the University of Leuven, where I presented earlier versions of these ideas, and Andrew Currie and two anonymous referees for *Acta Philosophica Fennica*, for their written comments on an earlier version of this paper.

References

Bykvist, K., and Hattiangadi, A. (2007), 'Does thought imply ought?', *Analysis* 67, pp. 277-285.

Chan, T. (ed.) (2013), *The Aim of Belief*, Oxford University Press, Oxford.

Chellas, B. (1980), *Modal Logic: An Introduction*, Cambridge University Press, Cambridge.

Charlow, N., and Chrisman, M. (eds.) (2016), *Deontic Modality*, Oxford University Press, Oxford.

Eikmeyer, H-J., and H. Rieser (eds.) (1981), *Words, Worlds, and Contexts: New Approaches in Word Semantics*, de Gruyter, Berlin.

von Fintel, K. (2001), 'Counterfactuals in a dynamic context', in M. Kenstowicz (2001), pp. 123-152.

Gillies, A. (2007), 'Counterfactual scorekeeping', *Linguistics and Philosophy*, 30, pp. 329-360.

Hughes, N. (ed.) (forthcoming), *Epistemic Dilemmas*. Oxford University Press, Oxford.

Kenstowicz, M. (ed.) (2001), *Ken Hale: A Life in Language*, MIT Press, Cambridge, MA.

Kratzer, A. (1977), 'What "must" and "can" must and can mean', *Linguistics and Philosophy*, 1, pp. 337-355.

Kratzer, A. (1981), 'The notional category of modality', in H-J Eikmeyer and H. Rieser (1981), pp. 38-74.

Kratzer, A. (2012), *Modals and Conditionals*, Oxford University Press, Oxford.

Matheson, J., and R. Vitz (eds.) (2014), *The Ethics of Belief: Individual and Social*, Oxford University Press, Oxford.

Moss, S. (2012), 'On the pragmatics of counterfactuals', *Noûs*, 46, pp. 561-586.

Portner, P. (2009), *Modality*, Oxford University Press, Oxford.

Williamson, T. (2020), *Suppose and Tell: The Semantics and Heuristics of Conditionals*, Oxford University Press, Oxford.

Williamson, T. (forthcoming), 'Epistemic ambivalence', in N. Hughes (forthcoming).

Worsnip, A. (2019), "'Ought'-contextualism beyond the parochial", *Philosophical Studies*, 176, pp. 3099-3119.

Guidance and Epistemic Filtering[1]

MARIA LASONEN-AARNIO

1. Guidance and access

Here is a familiar thought: the normative is capable of *guiding* us. If a theory, whether in ethics or epistemology, cannot generate guidance-giving norms, that is a mark against the theory. For instance, classical utilitarianism has been thought by many to fail disastrously on the guidance front.[2] Epistemologists have likewise invoked guidance, most often to criticize various externalist norms and theories.[3]

My starting point will be a rather ambitious view on which the normative is capable of guiding us invariably, in every possible situation. Such a thought has been attractive to many who think that guidance is necessary for acting as one ought in a way that is *creditworthy*, and thereby not merely lucky or accidental. We should always be able to act as we ought in a creditworthy way, the thought goes, which in turn requires an ambitious kind of guidance.

Such guidance ambitions encounter a grave problem. For almost everyone agrees that guidance requires some form of *epistemic access* to a domain of facts: in order to be guided by a normative principle, one must have access to whether the conditions specified by the principle apply, and in order to be guided by a normative reason, one must have access to the

---

[1] This article heavily draws on Lasonen-Aarnio (2019), generalizing its main line of argument.

[2] Jackson (1991) is a classic example.

[3] There is a vast amount of papers using guidance to argue for or against various epistemic norms. For a guidance-based argument against externalist norms, see e.g. Pollock (1987). For a guidance-based criticism of a truth norm on belief, see Glüer & Wikfors (2009, 2013). These are just some examples.

reason itself.[4] At first considerations having to do with guidance were seen by many as favouring more subjective theories, theories generating norms that make reference to a domain of facts within our ken.[5] But more recently sophisticated work in epistemology has questioned whether appeal to the subjective is any solution at all, at least given the ambitious view on which normative principles must always be capable of guiding us. Internalizing or subjectivising the facts grounding normative truths only yields unfailing guidance if we invariably have access to these facts. But we are sometimes in no position to access our beliefs, desires, appearances, seemings, motivations, or the quality of our will – indeed, there is no domain of facts that we can invariably access.[6] For this reason going internal or subjective does not guarantee guidance.[7]

This is the *Access Problem* for guidance. If there is no domain of facts we can invariably access, then it looks like there simply *are* no principles or norms that we are always in a position to be guided by in the desired way, though there might be a different, weaker kind of guidance that even more externalist or objectivist norms can provide. If the normative is essentially and invariably action-guiding, then we are forced to draw the grim conclusion that there is no such thing. If normativity is to be salvaged, we must, it seems, settle for less and tone down our guidance ambitions.

My focus here will be on views, gaining in popularity, that seem to evade the Access Problem altogether, and that there-

---

[4] For some representative examples, see Jackson (1991), Raz (2011, p. 110), and Gibbons (2013, p. 132).

[5] For instance, Hudson (1989) explicitly characterizes subjective theories in terms of guidance.

[6] Williamson (2000, Ch. 6) argues that there are no non-trivial *luminous* conditions, conditions such that whenever they obtain, we are in a position to know that they obtain. For a criticism, see Berker (2008). For a reply and clarification of the argument, see Srinivasan (2015b). For what I am calling the Access Problem, see Srinivasan (2015a) and Hughes (2018). See also e.g. Gibbons (2013, pp. 130-131). Worries about epistemic access in connection with guidance have been expressed by numerous moral philosophers. See, for instance, Smith (1988, 2010, 2012). See also Lasonen-Aarnio (*Forthcoming C*).

[7] See e.g. Srinivasan (2015a) and Hughes (2018).

by seem to rehabilitate an ambitious kind of guidance. Let us concede that the project of identifying some domain of facts that we always have access to is doomed to fail. But here is a thought: perhaps, instead, we can use epistemic access as a *criterion* on facts that determine normative facts in the first place, thereby requiring that such facts pass an *epistemic filter*.[8] Those who press the Access Problem can no longer complain that we don't always have access to the facts that ground normative truths, for access is now a requirement on being part of that domain! For instance, the most prominent defences of the Access Problem appeal to *anti-luminosity arguments*, which purport to show that there are no non-trivial *luminous* conditions, conditions such that whenever they obtain, we are in a position to know this. But if the facts that ground normative truths must pass a *knowledge filter*, then an anti-luminosity argument cannot be levelled: the challenge that we are sometimes in no position to know these facts simply cannot be raised.

Epistemic filtering is normally implemented in a reasons-centered picture of normativity. It is first assumed that normative facts hold in virtue of facts about normative reasons: if I ought to ϕ (in a relevant sense of 'ought'), that is so in virtue of the fact that the overall balance of my reasons supports ϕ'ing. But there is an epistemic filter on these reasons: a fact can only count as a normative reason for a subject if she has epistemic access to it.[9] According to a different implementation of filtering, only normative reasons that one *has* or *possesses* determine the relevant normative facts, and there is an epistemic access condition on possession.[10] On either view,

---

[8] This response goes back at least to H. A. Pritchard's discussion (see Pritchard 1932). Dancy (2000, p. 57) mentions the possibility of an "agent-relative epistemic filter". See also Raz's (2011, p. 110) discussion of epistemic filters.

[9] I take Markovits's (2010) view to be that only known facts can constitute right-making reasons. Gibbons (2013) defends a similar view of normative reasons. Similarly, according to Kiesewetter (2017), only propositions that are part of one's evidence can be reasons to begin with, and being known is at least sufficient for being part of one's evidence (e.g. p. 162, 200). See also Raz (2011, Ch. 6).

[10] According to Lord (2018, Ch. 3), in order to possess *r* as a reason (to ϕ), one must know *r*. This, Lord thinks, is not sufficient, for one might not

the *potent* normative reasons—reasons that enter into determining normative facts—are subject to an access-constraint.[11]

There are many ways of thinking about epistemic access, and correspondingly different kinds of epistemic filters. What makes for access to a fact? On one view, it is mere true belief. On another, it is justified or rational true belief. Perhaps the most popular (and to my mind plausible) view is that access is a matter of knowing. Indeed, knowledge has been the most commonly invoked epistemic filter: potent normative reasons must be known, or at least one must be in a position to know them.[12] But there are yet more options. Perhaps access requires a kind of apparentness, of something *seeming* to be the case: we have access to those facts that seem to us to be the case.

I will argue that, irrespective of the kind of filter deployed, epistemic filtering does not restore an ambitious kind of guidance. The reason for this is that a central good that guidance has been invoked to deliver is the availability of acting (or believing) as one ought in a way that is not merely lucky or accidental. But epistemic filtering does not guarantee the availability of such non-lucky ought-doing. In short, the problem is that given epistemic filtering, normative facts come to depend on *facts about what one has access to*, or *facts about what passes the epistemic filter*. But irrespective of the filter deployed, we are sometimes in no position to track these access-facts themselves. As a result, sometimes we can only act as we ought by luck.

2. Guidance and luck

Those who work within a reasons-centered picture of normativity tend to think of normative guidance as a matter of *responsiveness* to the normative reasons in virtue of which one

---

'see' the connection between *r* and ϕ'ing – I might, for instance, know that the fish contains salmonella but not see that this is a reason not to eat the fish if I falsely believe salmonella to be a harmless bacterium.

[11] I borrow the term "potent normative reason" from Littlejohn (2018).

[12] For two recent books defending versions of it, see Kiesewetter (2017) and Lord (2018).

ought to do this or that.[13] This responsiveness, in turn, is a matter of being motivated to act by those reasons. Thus, Way and Whiting (2017, p. 364) simply express the idea that the normative must be guiding as the familiar thought that "normative reasons can be motivating reasons".[14] We can always be guided by our reasons, then, just in case the following is true:

> *Guidance*
> If one ought to ϕ, then one can ϕ for the normative reasons in virtue of which one ought to ϕ.[15]

In what follows, my focus will be on views that endorse *Guidance* together with some form of epistemic filtering.

The discussion below will be focused on a central good that guidance has repeatedly been invoked to deliver. It is often assumed that guidance is necessary and sufficient for acting as one ought in a way that is not a mere accident, fluke, coincidence, or matter of good luck: being guided by a principle is often opposed with mere accidental conformity to it.[16] Indeed, a common complaint regarding objectivist theo-

---

[13] This view of guidance is in contrast to one on which guidance is a matter of using a normative principle in deliberation about what to do. For discussions of the deliberative view, see Smith (2012) and Hughes (2018); see also Pollock's (1987, p. 64) discussion of what he terms the "intellectualist model".

[14] See also Gibbons (2013, p. 135), Kiesewetter (2017, p. 11), Lord (2015, 2018).

[15] There are many ways to express similar ideas. For instance, Lord (2015) says that "the facts that obligate us must be potentially action-guiding", and that we must at least have the *ability* to act for the reasons determining what one is obligated to do. See Way & Whiting (2017) for a critical discussion of what it means to have the "ability" to act for the right reasons. See also Väyrynen (2006). It should be clear, however, that the 'can' here is not just that of metaphysical possibility.

[16] See, for instance, Kant's famous remark in *Groundwork* that "in the case of what is to be morally good it is not enough that it *conform* with the moral law but it must also be done *for the sake of the law*; without this, that conformity is only very contingent and precarious" (G 4:390). For more recent examples, Wedgwood (2002) says that being guided by a rule is incompatible with its being "purely a *fluke*" that one conforms to it. Smith (2012, p. 370) remarks that if one is not guided by a principle, one's conformity to it is purely "coincidental". Gibbons (2013, p. 128), discussing

ries and principles is that if these theories are true, and we lack access to the relevant domain of objective facts (as we often do), then sometimes we can only do what we ought in a haphazard, lucky or accidental kind of way. And many have found the thought that our normative lives are thus subject to fortune very difficult to live with.

But why is such non-lucky ought-doing important in the first place? Most parties agree that success that is not due to luck is closely connected with credit or praise. Doing what one ought not just by luck, or conforming to norms in a way that is not lucky or accidental, is necessary for conforming to them in a way that is *praiseworthy* or *creditworthy*. When one's action has *moral worth*, for instance, one is praiseworthy for doing the right thing.[17] A core motivation for guidance, then, is a need, via the notion of luck, to tie together the *deontic* and *hypological*.[18] The thought is that it should always be possible to act as one ought in a *creditworthy* or *praiseworthy* way—and hence, in a way that is not merely lucky. *Guidance* is supposed to guarantee this, in so far as non-lucky, creditworthy ought-doing just is a matter of ϕ'ing for the potent normative reasons in virtue of which one ought to ϕ.

---

objectivist norms, complains that "if we did manage to comply with the objectivist's norm, that would just be an accident". Väyrynen (2006) discusses the idea that without guidance, a successful moral life would only be available "by luck or happenstance". This is just a sample: it is difficult to find any discussion of guidance that does *not* draw on the idea that without guidance, one's conformity to a principle or norm is merely accidental.

[17] Doing the morally right thing not by mere luck or accident plays a pivotal role in discussions of morally worthy action. Most authors talk about morally praiseworthy action (e.g. Markovits 2010, Arpaly 2003, Sliwa 2015). Johnson King (2020) distinguishes between different types of praiseworthiness, arguing that when performing a morally worthy action, one is praiseworthy *for acting rightly*. I agree with Johnson King that one might be praiseworthy for doing something even if the action is morally right only in an accidental way.

[18] For this terminology, see Zimmerman (2002, p. 554); see also Srinivasan (2015a). For instance, Lord (2015) thinks that non-accidentality is required for creditworthiness, and that it would be a big cost to admit the possibility of cases in which one cannot act in a creditworthy way.

In what follows, I will assume that an adequacy condition on views of guidance is that being guided in the desired way entails that one acts as one ought in a way that is not merely lucky. I will argue that a responsiveness view of guidance, even when coupled with epistemic filtering, does not guarantee the availability of non-lucky ought-doing: either one cannot always be motivated to act by one's potent normative reasons, or acting for the potent normative reasons in virtue of which one ought to act does not rule out its being lucky or accidental that one acted as one ought.

My argument will draw on the idea that a certain kind of modal robustness or invariance of a success is *necessary* for its not being merely lucky. As I see it, such a modal condition is rather weak, and compatible with various full accounts of luck. Here are the bare bones of the modal condition I have in mind. Assume that a subject ϕ's—chooses a particular course of action, performs that action, forms a belief, etc. Assume that her ϕ'ing is a normative success: she chooses, acts, or believes as she ought, as a relevant norm tells her to. In order for her success to not be lucky, her *way* of ϕ'ing must issue in normative success in a sufficiently invariant manner across a range of *relevant* counterfactual cases in which she chooses, acts, or forms a belief in that way. (The reader should not assume that such modal robustness is to be understood in terms of modally *close* cases.)[19] ϕ'ing in a way that is robustly successful in this sense is necessary for one's success to not be merely lucky. Think of the relevant cases, very roughly, as ones in which we allow the idiosyncratic features of the actual case to vary in acceptable, somewhat normal ways. The idea is that if one's success depends on the obtaining of some idiosyncratic feature of one's situation, then it is merely lucky.

The epistemology literature is replete with proposals for how to understand evaluations sensitive to, in particular, ways of forming beliefs: a belief is formed in a good way if it

---

[19] Cases that are modally close, involving events that could easily have occurred, are often relevant. Ultimately, however, I don't think relevance is a matter of any kind of similarity relation (for more details, see Lasonen-Aarnio *forthcoming B, forthcoming C*). Hence, I am *not* assuming that praiseworthiness or creditworthiness requires that a success be *safe*.

is properly based on sufficiently good, undefeated *reasons* or *evidence*, if it is the output of a *reliable process*, or it is formed by a reliable *method*. Ultimately, I prefer to think of these ways in terms of the dispositions that manifest as one's ϕ'ing, though the main points I make below are compatible with alternative views.[20] The general proposal, then, is that in order for one's ϕ'ing to be an instance of non-lucky ought-doing, ϕ'ing in that way must issue in normative success in a sufficiently invariant way across a relevant range of counterfactual cases. One's way of ϕ'ing must track what one ought to do across the relevant cases.[21] Given a dispositional understanding of ways, non-lucky ought-doing requires manifesting dispositions that don't, across a range of normal cases, manifest as acting in ways that one ought not to act.[22] If I am in no position to manifest dispositions that *discriminate* between relevant cases in which I ought to ϕ, and those in which I ought not to ϕ, then I am only in a position to act as I

[20] My reasons for adopting the dispositional view have to do with its theoretical utility. Understanding ways in terms of dispositions allows giving a unified account of non-lucky successes—whether ones involving conscious deliberation, automatic actions, doxastic revision, or athletic performances. Elsewhere I argue that it enables making sense of the verdict that victims of massive deceit can be forming their beliefs in ways that are just as good as those of ordinary subjects—thereby meriting a kind of praise—and that subjects who obstinately stick to their beliefs in putative cases of higher-order defeat are criticizable for managing their beliefs in bad ways. See Lasonen-Aarnio (*forthcoming A*, *forthcoming B*).

[21] Tracking in this sense is not trivial even if I ought to perform the same action, or form the same belief, across all of the relevant cases. Assume, for instance, that we want to evaluate whether my way of forming a belief about the result of adding 126 to 296 tracks the truth of the matter. Even if the answer is the same across all cases, tracking the correct answer is not trivial: randomly guessing or adding up numbers using a defective method don't track the relevant mathematical fact across relevant cases.

[22] Note that on my view it does not matter if the fact that one manifests the disposition in question is modally fragile. It might, for instance, be a very local and circumscribed disposition, even one that is only acquired in specific circumstances (see Lasonen-Aarnio *forthcoming B*). Thanks to an anonymous referee for bringing this up.

ought by a kind of luck. This sort of *dispositional discrimination* will play an important role in the discussion below.[23]

It is helpful to see how the modal condition diagnoses why one is sometimes in no position to conform to objectivist norms or principles save by luck. Consider a well-known case from Jackson (1991, p. 463) involving *Dr Jill*, a physician who must choose which of three drugs to prescribe for her patient with a skin condition. Jill knows that drug A is likely to relieve the patient's condition without completely curing it. Further, Jill knows that one of drugs B and C will completely cure the condition, while the other will kill the patient. However, she has no way of telling which is which. Appeal to such examples has become the standard argument against objectivist theories and norms, such as a norm telling one to do what is best. In this case, this norm would require prescribing whichever of B and C cures the condition. But Jill is in no position to prescribe the best drug except by good luck! The modal account provides the following diagnosis of why this is so: no way of choosing between B and C is available to Jill that would track what is best across a range of counterfactual cases. Jill could, for instance, choose randomly, but randomly choosing would lead her to prescribe the wrong drug across a range of cases.

At first sight, it looks like a view endorsing epistemic filtering, filtering of potent normative reasons by some type of epistemic access, is not susceptible to the challenge that one can sometimes only act as one ought by luck. In what follows, I will argue that epistemic filtering does not solve the problem of luck, for it makes normative facts depend on facts about just what propositions we have epistemic access to, *and we are sometimes in no position to track these facts*. I first discuss knowledge filtering. I then generalize my arguments to other kinds of epistemic filters.

---

[23] For more on this notion of discrimination, and how it differs from an *epistemic* kind of discrimination, see Lasonen-Aarnio (*forthcoming B, forthcoming C*).

3. Knowledge filtering: a case study[24]

Knowledge is the most commonly invoked epistemic filter: potent normative reasons must be known, or at least one must be in a position to know them.[25] Recall that according to the view of guidance under focus, the *responsiveness* view, potent normative reasons must be capable of serving as our motivating reasons. There is an impressive array of arguments for the conclusion that something can be a motivating reason only if it is known.[26] Further, guidance is often taken to entail acting as one ought in a way that is not merely accidental or lucky, and several authors have argued that no epistemic standing short of knowledge can guarantee such non-accidentality.[27]

Consider, then, the following view:

Knowledge Filtering
p is a potent normative reason for *s* just in case *s* knows *p*.[28]

I will argue that knowledge filtering does not solve the guidance problem, for it does not guarantee the availability of acting as one ought in a way that is not lucky, a central good that guidance has been invoked by many to deliver. Interestingly, the reasons for this derive from some of the very same structural features of knowledge that the Access Problem rests on.

Knowledge Filtering guarantees that ought-making facts are never beyond one's epistemic ken. But one of its immediate consequences is that normative facts now depend on epistem-

---

[24] This section draws very heavily on Lasonen-Aarnio (2019).

[25] For two recent books defending versions of it, see Kiesewetter (2017) and Lord (2018); see also the references in note viii.

[26] The view goes back at least to Unger (1975), and has been more recently defended by John Hyman, Timothy Williamson, and Jennifer Hornsby. For references, see Alvarez (2016).

[27] See, for instance, Sliwa (2015).

[28] Instead of being known, one might think that it suffices that *s* be in a position to know *p*. The discussion below applies equally to such views. Lord thinks knowledge is not sufficient (see note 9). However, his reasons don't have to do with the problems I raise below. As a result, as far as I can see, the extra condition he proposes on one's possessing *r* as a reason to ϕ doesn't solve these problems.

ic facts, facts about what one knows, or is in a position to know. Indeed, this is one of the main selling points of views that appeal to epistemic filtering, often defended by considering Jackson-type cases.[29] In the original *Dr Jill* case, it is false that Jill ought to prescribe drug B. But now assume that she learns that drug B is in fact the cure, and C the killer. Given her new epistemic position, she ought to prescribe drug B. The only relevant change is her coming to know which drug cures the skin condition. True, the normative reason Jill acquires is *that drug B is the cure*, and not *that I know that drug B is the cure*: we can distinguish between what is in the set of potent normative reasons and criteria for getting there. But there is at least a modal, counterfactual kind of dependence between facts about one's epistemic position and normative facts. This point is worth emphasizing, for one of its consequences, as I will argue, is that sometimes non-lucky ought-doing will require being able to track the relevant facts about one's epistemic position in a modally robust way. And this, given the nature of knowledge, is not something we can always do.

I will now describe two kinds of problem cases, given in Lasonen-Aarnio (2019), in which the relevant subject is in no position to track those facts about what she knows that a relevant normative fact depends on and hence, *assuming knowledge filtering*, in no position to act as she ought in a way that is not lucky. These cases follow two kinds of templates: In the first, the normative fact that one ought to ϕ essentially depends on one's knowing (or being in a position to know) a relevant proposition *p*, which is a reason to ϕ: given that this proposition is among one's normative reasons, one ought to ϕ. But all else being equal, if this knowledge condition was not met and hence, if *p* was not among one's normative reasons, it would be false that one ought to ϕ. The second kind of template is one in which the normative fact that one ought to ϕ essentially depends on one's *not* knowing a relevant proposition *p*.

[29] E.g. Kiesewetter (2015, 2017) and Lord (2015, 2018).

Margins for error and precarious knowing

Assuming that knowledge is subject to a margin for error principle, knowing can sometimes be precarious in the sense that one could easily have been in no position to know a relevant proposition.

Li works Saturday mornings at a climbing gym, when the place is flooded by young children and their parents taking over the auto belay devices. Li has to frequently estimate the weight of a child just by looking, for the scale at the gym is highly unreliable, and it is important that children too light are not allowed to use the auto belay devices. Li has become very good at estimating how much a child weighs just by looking: he can normally tell by a margin of 1 kg. Now consider the following example.

> *Auto belay*
> Li knows that a child weighing exactly 15 kg or more is heavy enough to be lowered down by one of the auto belay devices. It would be a disaster, however, to let a child lighter than 15 kg climb: if they eventually got up, they would be left dangling, probably screaming, on the wall 16 meters above the ground, and it would be difficult to get them down – not to mention a *huge* PR disaster for the gym. Li sees a certain high-profile parent involved in local politics walk in with their daughter Ada. There is a lot of pressure to let the child climb, for the whole future of the gym depends on a favour from the parent. In fact, Ada weighs 16,05 kg. Li comes to know, just by looking, that she weighs over 15 kg.

I will assume a margin for error principle on knowledge. In particular, given the accuracy of Li's estimates, had Ada been just 100 grams lighter, Li would have been in no position to know that she weighs at least 15 kg. As a result, Li *just barely* knows that Ada weighs at least 15 kg. In this sense, Li's knowledge is precarious: he could easily not have known. Moreover, given the absence of a sufficiently reliable scale (or parent), he could easily have been in no position to know.

I will make the following assumptions about the case: *Ada weighs 15 kg or more* is an excellent reason to let her climb. Hence, given *Knowledge Filtering*, as long as Li knows this, he ought to let her climb. Moreover, it is difficult to see why the

proposition that Ada weighs at least 15 kg couldn't be his motivating reason for letting her climb. It is, however, essential that this proposition be among his reasons, given that it would be very bad to let a child weighing less than 15 kg climb. Hence, propositions like *it is probable that Ada weighs 15 kg or more*, or *it seems that Ada weighs 15kg or more* just wouldn't cut it. (I revisit this below.) As a consequence, had she been 100 grams lighter, the normative facts would have been different due to Li's being in no position to know that the child weighs at least 15 kg.

Li's knowledge is precarious. He could very easily have not known, and no way of acting or choosing is available to him that discriminates between his case and the case in which Ada is just 100 grams lighter. Moreover, this case is relevant for evaluating whether Li's normative success is lucky: being in no position to track the normative facts given very slight changes to the child's weight means that Li is in no position to act as he ought in a manner that is not subject to luck.[30] Similarly, consider a counterfactual case in which the child is a bit lighter, weighing 15,95 kg. In that case Li ought not to let the child climb, since *Ada weighs 15 kg or more* is not among his potent reasons. But again, Li is in no position to act as he ought in a way that is not lucky, since he could easily have known, in which case the normative facts would have been different.

One might, of course, quibble about the details of the case. But if knowing can be precarious in the way assumed, and if a relevant normative fact can essentially depend on one's having such precarious knowledge, then examples with the structure described are bound to arise.

Those who resist anti-luminosity arguments might argue that the kind of precarious knowing I have described does not exist. But at this point it is worth reminding the reader of the dialectic. My focus has been on appeal to epistemic filtering as an answer to the *Access Problem* for guidance. In a nut-

---

[30] Note that I am *not* claiming that, if Li lets the child climb, it is accidental that the child gets down safely – arguably, a margin for error principle on knowledge guarantees that *this* success is not accidental. But that is not the question: the question is whether, assuming *Knowledge Filtering*, the success of *acting as he ought* is subject to luck.

shell, the problem is that since there is no domain of luminous facts, guidance ambitions are doomed to fail. Epistemic filtering, it seemed, offered a way to sidestep the problem without contesting anti-luminosity reasoning. I have argued, in effect, that the very structural facts about knowledge that anti-luminosity arguments rely on create a problem of luck even assuming *Knowledge Filtering*.

The second example does not rely on the assumption that knowledge is subject to a margin for error principle.

Unlucky not to know

Again, Jill must decide which of three drugs to prescribe for her patient with a skin condition. Several years ago Jill carefully researched the matter. Jill rates drugs using a scale from 1 to 10, 1 being maximally harmful (killing a patient), 10 being maximally helpful (a complete cure), and 5 being neutral. She remembers rating A as a 7. She remembers assigning either B or C a rating of 10, while assigning the other a 1, but she cannot remember which is which! Thankfully, she wrote the conclusions of her research in notebook #158, which is in her office. There is just enough time to consult the notebook. Now consider:

> *Notebook trouble*
> Jill rushes to her office, and pulls out notebook #158. She opens the page on which she compiled her main findings on drug C, and written in the bottom of the page is the number 1: C is the killer drug, and therefore, Jill concludes, B is the cure! Alas, though this is true, Jill does not know it. About a week ago Jill had to leave her children alone in her office for 5 minutes. Four-year old Fanny, looking for something to apply her new eraser to, happened to pull out notebook #158 from the bookshelf and started randomly erasing numbers. She soon ended up on the page describing drug C, carefully erasing the '1' Jill had written on the bottom of the page. At this point eight-year old Eartha realized what was happening. Eartha asked what had been written on the page. The number 10, Fanny confidently (but falsely) told her. Not wanting to get into trouble, Eartha began carefully writing '10' on the bottom of the page. Very soon after she began, before she had time to write '0', she heard her mother's

> footsteps in the hallway, quickly closed the notebook, and put it back in its place.

Dr Jill holds a justified, true belief that drug B will cure the skin condition, but lacks knowledge, for she is in a Gettier case: she could very easily have falsely believed that C is the cure, and B the killer, which she would have done had Eartha had time to complete her task. In fact, she may currently not even be in a position to know which drug is the killer. Hence, assuming knowledge filtering, the proposition *drug B will cure the skin condition* is not among Dr Jill's potent normative reasons. Jill may know that it is likely on her evidence that drug B is the cure and that drug C the killer, but assuming the consequences of prescribing the wrong drug to be harmful enough, this does not suffice to make it the case that she ought to prescribe B. Similarly, she might know that it appears to her that drug B is the cure. But a mere appearance is compatible with B being harmful, and does not weight heavily enough in favour of prescribing B. In the absence of the proposition *drug B is the cure,* Jill's potent normative reasons just don't weight heavily enough in favour of prescribing drug B. (Below I discuss a reply insisting that they do.)

The totality of potent normative reasons that Jill has in *Notebook trouble* is not relevantly different from those in our original *Dr Jill* case. Since she doesn't know that drug B is the cure, what Dr Jill ought to do is prescribe medicine A (that is the standard verdict in the original case).[31] The important question is: can she act as she ought in a way that is not merely lucky or accidental? I think not. Relatedly, if Jill ends up prescribing drug A, acting as she ought, her normative success is not creditworthy or praiseworthy. This verdict can be supported by appeal to the modal condition on luck. Jill cannot discriminate her case from counterfactual cases in which she is not in a Gettier case, and does know that B is the cure on the basis of consulting her notebook. Indeed, had everything been normal, she would have known. Any way of making a choice that results in prescribing medicine A will lead her astray in cases in which she knows that B is the cure.

---

[31] Note that nothing I have said explains how (or whether) proponents of *Knowlege Filtering* get the verdict they want in our original case of *Dr Jill*, but I am setting this problem aside.

Compare a subject in a Gettier case and a subject who forms a true belief on the basis of a random guess. Neither subject knows. However, as several epistemologists have pointed out, only the subject in a Gettier case is *unlucky not to know*.[32] Jill is unlucky not to acquire knowledge by consulting her notebook. The modal condition on luck outlined above explains this, for the way in which she forms her belief results in coming to know across a wide range of relevant, somewhat normal cases. As a result, her failure to know is merely unlucky.

One way to try to block my argument is to concede the claims I have made about knowledge, but to contest their implications for normativity in the context of *Knowledge Filtering*. In particular, one could contest the assumed dependence between facts about what Li and Jill know and facts about what they ought to do. Even had Li not known that Ada weighs at least 15 kg, its *seeming* to him that way would have sufficed to make it the case that he ought to let her climb. And in *Notebook trouble* Jill ought to prescribe medicine B after all, for even though she does not know that it is the cure, it *seems* or *appears* to her to be the cure, and she knows which number is written in her notebook, and these reasons are sufficiently strong on their own.

Instead of debating the details of particular examples, we should ask what must be assumed for such a reply to be available across the board. One would have to subscribe to the general claim that *p* can't be a weightier reason than an appearance or seeming that *p*: if *p* would (if known) be a potent normative reason for *s* to ϕ, then *its seeming to one that p* (or the seeming itself) would be a potent normative reason to ϕ with the same weight.[33] Moreover, in the overall balance of reasons, *its seeming that p* as it were screens off whatever weight *p* itself has: if, for instance, it first merely seems to me

---

[32] See Sutton (2005) for a similar point about Gettier cases. Sutton thinks this is true, more generally, of subjects who hold justified beliefs that fail to constitute knowledge. Hirvelä (MS) argues that subjects in Gettier cases could *easily* have known.

[33] Kiesewetter (2017, Ch. 7) endorses a version of the backup view in response to an objection to his view of rationality. By contrast, Lord (2018, p. 193) concedes that what he calls the 'atomic' weight of a reason like *drug B is the cure* can be greater than that of *it appears that drug B is the cure*.

that *p*, and I then come to (be in a position to) know *p*, acquiring the new potent normative reason that *p* can make no difference for what I ought to do. Among our potent normative reasons, propositions concerning the world outside our appearances turn out to be normatively inert! I think this is extremely implausible. For instance, if *p* (obviously) entails *q*, then *p* is a maximally strong reason to believe *q*. But its seeming to me that *p* is not a maximally strong reason to believe that *q*, for it is compatible with *q* being false. One would think that the strength of one's reasons for believing *p* had something to do with probabilities. But unless one's prior probabilities *rule out* the possibility of misleading seemings or appearances, the probability of *q* on *p* will be higher than the probability of *q* on its seeming to one that *p*.[34] I deliberately chose the examples so that the consequences of acting if the relevant propositions are false are disastrous, and in such cases it matters whether one's reason is that it seems that *p*, or that *p* is the case.

But ultimately, even the backup view does not solve the problem. Let me once again remind the reader of the dialectic here. Anti-luminosity arguments, it seemed, simply have no bearing on views that have an epistemic filter on the potent normative reasons. But whether appeal to appearances or seemings, together with the backup view discussed, solves the problem I have raised depends on what it takes for facts about seemings to become potent normative reasons. If a proposition must be known in order to be among one's potent reasons, then the proposed view is susceptible to anti-luminosity type reasoning: there are cases in which it seems to one that *p*, but one is in no position to know this.

It is also instructive to see why piling on more access conditions won't help. Consider a more stringent epistemic filter: in order for a proposition *p* to pass the filter, one must know that one knows *p*. Normative facts now depend on facts about whether one has such iterated knowledge. The problem is that there will be cases in which *p* is an essential normative reason weighting in favour of ϕ'ing, but one is in no position to track the relevant epistemic facts, facts regarding whether one knows that one knows *p*. That is because knowledge that

[34] See also Littlejohn's (2018) discussion of the backup view.

one knows *p* is susceptible to exactly the same phenomena as knowledge that *p*: one's knowledge that one knows can be precarious—one could easily have merely known—and one can have a Gettier belief that one knows.

At this point those who appeal to seemings might simply reject knowledge filtering. True, they might concede, seemings are not immune to anti-luminosity reasoning. But we should simply reject knowledge as the relevant kind of epistemic access. Guidance requires access, but the access in question itself is a kind of apparentness: we have access to just those facts that *seem* to us to be the case. Alternatively, perhaps seeming-states *themselves* figure among our potent normative reasons, and do so *as long as we undergo them*.[35] How is anti-luminosity type reasoning relevant against such views in the first place? In the next section my aim will be to generalize the problem to this alternative picture of epistemic access.

4. Seemings filtering: a second case study

Some might worry that structural analogues of the anti-luminosity problem kept cropping up because of a knowledge-centric view of epistemic access. So it is worth discussing a wholly different kind of epistemic filtering.

Here is a thought. Only some facts are made apparent to us. Being made apparent in the relevant sense is not a matter of being known. It is a matter of these facts *seeming* to us to be the case, where a seeming is a *sui generis* kind of conscious mental state with propositional content and a distinctive phenomenology.[36] To be among our potent normative reasons, facts have to pass a *seemings filter*: in order for the fact that *p* to be among my potent normative reasons (or among the facts grounding the relevant normative truths), it must seem to me that *p*.

---

[35] I take this to be Kiesewetter's (2018) view, though his views is not that *only* such seeming-states constitute potent normative reasons.

[36] This phenomenology has been described as 'forceful' (Huemer 2001, pp. 77-9), 'assertive' (Tucker 2010, p. 530) as 'recommending' its content, and as having 'the feel of truth' (Tollhurst 1998, p. 298). As Siegel pointed out to me, the idea goes back at least to Price's 1932 book *Perception*. See further references in note 13 of Siegel (2017).

Here is another, slightly different thought. Ultimately, (potent) normative reasons are not facts or propositions, but mental states. And they are, in particular, *seeming-states*. Whether or not it is true that *p*, when it seems to me that *p*, the seeming-state itself is a potent normative reason, the totality of my potent reasons consisting in the totality of such seeming-states.[37] All it takes for a seeming-state to be among my potent normative reasons is for me to *undergo* that state.

These views differ in important respects. Assume that it seems to me that the pond is frozen, but it is in fact not frozen. Since it is not a fact that the pond is frozen, according to the first view there is no work for a seemings filter to do: at least the proposition that the pond is frozen is not among my potent normative reasons. According to the second view, whether or not the pond is frozen makes no difference, for either way, the seeming-state itself, with the content that the pond is frozen, is among my potent normative reasons. But despite their differences, according to both views, just what my potent normative reasons are – and hence, what the relevant normative facts are – depends on what seeming-states I undergo. That is, normative facts come to depend on facts about my seemings.

I argued above, drawing on previous work, that we are sometimes in no position to dispositionally discriminate between cases in which we know and cases in which we don't: sometimes we just cannot track the relevant knowledge facts and hence, the normative facts that depend on them. As a result, sometimes we are in no position to act as we ought in a way that is not lucky. Might seemings filtering immunize one from this sort of worry? A case in which one knows *p*, and a case in which one does not know *p*, can as it were look just the same from the inside – indeed, those who deploy seemings-talk will be inclined to say that things can *seem* just the same, for knowledge depends on facts about the world and one's relation to it. This was clearest in the second kind of case discussed above, which was a Gettier case. By con-

---

[37] According to a hybrid view, both mental states such as seemings, and propositions (presumably passing some epistemic filter) can be among one's potent normative reasons. I won't separately discuss such hybrid views, for the points I make equally apply to them.

trast, if it seems to a subject that p, a case in which it does not so seem cannot, trivially, be one in which things seem exactly the same!

But essentially the same problem, I want to argue, arises given seemings-filtering. Indeed, it was not even assumed in the precarious knowing case that Li's phenomenology is exactly the same in a case in which he knows and a case in which he does not. Even if seemings with different contents differ in some way regarding their phenomenology, substantive argument would be needed to establish that the possibilities regarding their contents always pattern with the limits of dispositional discrimination for a subject – that the limits of dispositional discrimination place principled limits on phenomenology. One would have to argue that the following kind of situation is impossible: while both a seeming that p_1 and a seeming that p_2 are in a subject's repertoire of possible experiences, she cannot discriminate between them. (Such dispositional discrimination would consist in reliably responding to these seemings in different ways across a range of relevant counterfactual cases – for instance, by coming to believe p_1 – not p_2 – when it seems to one that p_1, and by coming to believe that p_2 – not p_1 – when it seems to one that p_2.)

But why think that phenomenology is thus constrained by one's abilities of discrimination? Consider, for instance, a seeming that an object is red_{27}, and a series of cases by which a subject very gradually comes to experience a seeming that it is some quite different shade of red, say red_{64} – perhaps, for instance, the light shining on the object very gradually and smoothly changes colour. Why couldn't her colour seemings change so gradually that she cannot dispositionally discriminate differences in how things seem from one moment to another, being unable to reliably track differences between very similar seemings across a range of relevant counterfactual cases? Once again precisely the kinds of considerations that anti-luminosity arguments rely on raise their head. Those who press the anti-luminosity argument will argue that seemings are not luminous: one is not always in a position to know how things seem. That, of course, is not what is at issue here. But underlying the plausibility of anti-luminosity reasoning is the assumption that it is just not feasible to be disposed to discriminate between two cases that are

phenomenologically very alike – indeed, I think that lack of such dispositional discrimination *explains* lack of epistemic discrimination, explaining why one is not always in a position to know exactly how things seem.

If I am right, then we can press cases against seemings filtering that are structurally similar to the margin for error cases discussed above in connection with knowledge filtering. Consider, for instance, the first kind of view outlined above: one's potent normative reasons consist in those facts that seem to one to be the case. We can now build cases according to the following recipe. First, some proposition p_1 is true, and in fact seems to one to be true. (p_1 could, for instance, be the proposition that a child weighs more than 15 kg, or the proposition that a certain object is shade red_{27}.) Hence, p_1 is among one's potent normative reasons. Second, p_1 is not only a reason for one to ϕ, but an essential one, in the following sense: if p_1 were removed from one's stock of potent normative reasons, it would no longer be the case that one ought to ϕ. Finally, one cannot dispositionally discriminate between a seeming that p_1, and various other seemings that one undergoes in other relevant counterfactual cases, such as a seeming that p_2 (that the child weighs just 15 kg, and no more; that the object is red_{28} instead of red_{27}, etc.), where the contents of these other seemings, even if true, are *not* reasons to ϕ. That is, it is not feasible for one to be disposed to respond differently to these different seemings. As a result, one cannot discriminate one's actual case, in which one ought to ϕ, from other relevant cases in which one ought not to ϕ. But then, if one ought to ϕ and does so, it seems that one is lucky to have acted as one ought.

The worry is reminiscent of some discussions of the so-called *problem of the speckled hen*: one's phenomenology can have a determinacy that outstrips one's discriminative abilities.[38] Some have appealed to a distinction between *seemings* and *experiences* in order to solve the problem of the speckled

[38] See, for instance, Pace's (2010) discussion, who recaps the core of the problem as follows: "there are some properties presented in experience at a level of determinacy that outstrips the subject's recognitional capacities" (Pace 2010, p. 404).

hen.[39] But a principled case would have to be made that while the contents of experiences can outstrip our abilities of discrimination, this could not happen for seemings. And as far as I can see, no such case has been made.[40] It is also worth noting that various strategies for dealing with such discrimination and access worries have no bite in the current dialectical context. For instance, in response to anti-luminosity worries applied to evidence – which he thinks consists in non-factive mental states – Declan Smithies (2012) invokes the distinction between propositional and doxastic justification: whenever I am in a given kind of mental state, I have propositional justification to believe that I am in that state, even if, due to my limited doxastic capacities, I cannot exploit that justification to come to justifiably believe that I am in that state. Note that this strategy consists in conceding my points about lack of dispositional discrimination, while insisting that even if I cannot dispositionally discriminate a seeming that the object is red_{27} and a seeming that it is red_{28}, I nevertheless have propositional justification to believe truths about how things seem. However, in conceding the point about dispositional discrimination, this strategy does nothing to help with the luck-related worry: when the fact that I ought to ϕ depends on its seeming to me that the object is red_{27} (and not that it is red_{28}), my lack of discrimination prevents me from acting as I ought in a way that is not lucky.

5. Conclusions: trouble for epistemic filtering

Normative guidance requires having access to facts in virtue of which normative facts hold. Think of these ought-makers as going into a box, the contents of which determine the relevant normative facts. The *Access Problem* arises because given a wide range of views about such ought-making facts, we *don't* always have epistemic access to them, not even if they belong to an internal domain concerning our beliefs, experiences, or motivations. Epistemic filtering would seem to bypass the problem, for it *guarantees* epistemic access.

---

[39] See, for instance, Tucker (2010), Brogaard (2013).

[40] These issues were discussed in an earlier, longer version of Hawthorne & Lasonen-Aarnio (*forthcoming*).

The new problem, however, is that epistemic filtering of any kind makes normative facts dependent on facts about what we bear the relevant epistemic relation to. But for any candidate kind of epistemic access, we are sometimes in no position to discriminate just which propositions we can access. As a result, we sometimes cannot act as we ought save by luck—and hence, we cannot act as we ought in a creditworthy way.[41]

University of Helsinki

References

Alvarez, Maria (2016), "Reasons for Action: Justification, Motivation, Explanation", in *Stanford Encyclopedia of Philosophy*.

Arpaly, Nomy (2003), *Unprincipled Virtue: An Inquiry into Moral Agency*, Oxford University Press.

Berker, Selim (2008), "Luminosity Regained", *Philosophers' Imprint* 8.2, pp. 1–22.

Brogaard, Berit (2013), "Phenomenal Seemings and Sensible Dogmatism", in Tucker (ed.) *Seemings and Justification: New Essays on Dogmatism and Phenomenal Conservatism*, Oxford University Press, Oxford, pp. 270-289.

Dancy, Jonathan (2000), *Practical Reality*, Oxford: Oxford University Press.

Gibbons John (2013), *The Norm of Belief*, Oxford: Oxford University Press.

Glüer, Kathrin & Wikfors, Åsa (2009), "Against Content Normativity", *Mind* 118.469, pp. 31–70.

Glüer, Kathrin & Wikfors, Åsa (2013), "Against Belief Normativity", in Chan (ed.), *The Aim of Belief*, Oxford University Press, pp. 80–99.

Johnson King, Zoë (2020), "Accidentally Doing the Right Thing", *Philosophy and Phenomenological Research* 100.1, pp. 186–206.

Hawthorne, John & Lasonen-Aarnio, Maria (Forthcoming), "Not So Phenomenal", *The Philosophical Review*.

Hirvelä, Jaakko, *MS* "Justification and the Knowledge Connection".

Hudson, James L. (1989), "Subjectivization in Ethics", *American Philosophical Quarterly* 26.3, pp. 221–229.

[41] Many thanks to Daniel Drucker, Giada Frantantonio, Jaakko Hirvelä, Antti Kauppinen, Lisa Miracchi, and Niall Paterson for discussions and comments on an earlier draft, and to comments from two anonymous reviewers. This research has been funded by a grant from the *European Research Council*.

Huemer, Michael (2001), *Skepticism and the Veil of Perception*, Rowman & Littlefield Publishers, Inc., Lanham.

Hughes, Nick (2018), "Luminosity Failure, Normative Guidance and the Principle 'Ought-Implies-Can'", *Utilitas*, 30.4, pp. 439–457.

Jackson, Frank (1991), "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection", *Ethics* 101.3, pp. 461–482.

Kagan, Shelly (2018), "The Paradox of Methods", *Politics, Philosophy & Economics*, 17.2, pp. 148–168.

Kiesewetter, Benjamin (2011), "'Ought' and the Perspective or the Agent", *Journal of Ethics & Social Philosophy* 5.3, pp. 1–24.

Kiesewetter, Benjamin (2017), *The Normativity of Rationality*, Oxford: Oxford University Press.

Lasonen-Aarnio, Maria (2010), "Unreasonable Knowledge", *Philosophical Perspectives* 24(1), pp. 1–21.

Lasonen-Aarnio, Maria (2019), "Guidance, epistemic filters, and non-accidental ought-doing", *Philosophical Issues* 29.1, pp. 172–183.

Lasonen-Aarnio, Maria (*Forthcoming A*), "Competent Failure and Victims of Deceit", in Dutant and Dorsch (ed.), *The New Evil Demon,* Oxford: Oxford University Press.

Lasonen-Aarnio, Maria (*Forthcoming B*), "Dispositional Evaluations and Defeat", in J. Brown and M. Simion (eds.), *Reasons, Justification and Defeat*, Oxford University Press.

Lasonen-Aarnio, Maria (*Forthcoming C*), "Perspectives and Good Dispositions", *Philosophy and Phenomenological Research*.

Lord, Errol (2015), "Acting for the Right Reasons, Abilities, and Obligation", Shafer-Landau R (ed.). *Oxford Studies in Metaethics*, Vol 10, Oxford University Press, Oxford, pp. 26–52.

Lord, Errol (2018), *The Importance of Being Rational*, Oxford: Oxford University Press.

Markovits, Julia (2010), "Acting for the Right Reasons", *Philosophical Review*, 119. 2, pp. 201–242.

Pace, Michael (2010), "Foundationally Justified Perceptual Beliefs and the Problem of the Speckled Hen", *Pacific Philosophical Quarterly* 91, pp. 401–441.

Parfit, Derek (1988), "What We Together Do", *Unpublished manuscript*.

Pollock, John (1987), "Epistemic Norms", *Synthese* 71.1, pp. 61–95.

Pritchard. H. A. (1932), "Duty and Ignorance of Fact", reprinted in *Moral Writings*, McAdam, Jim (ed.), *Clarendon Press*, *Oxford,* 2002, pp. 84–101.

Raz, Joseph (2011), *From Normativity to Responsibility*, Oxford University Press, Oxford.

Siegel, Susanna (2017), *The Rationality of Perception*, Oxford: Oxford University Press.
Smith, Holly M. (1988), "Making Moral Decisions", *Noûs*, 22.1, pp. 89-108.
Smith, Holly M. (2010), "Subjective Rightness", *Social Philosophy and Policy* 27.2, pp. 64–110.
Smith, Holly M. (2012), "Using Moral Principles to Guide Decisions", *Philosophical Issues* 22, pp. 369–86.
Smithies, Declan (2012), "Mentalism and Epistemic Transparency", *Australasian Journal of Philosophy*, 90:4, pp. 723-741.
Srinivasan, Amia (2015a) "Normativity Without Cartesian Privilege", *Philosophical Issues* 25-1, pp. 273–299.
Srinivasan, Amia (2015b), "Are We Luminous?", *Philosophy and Phenomenological Research* 90.2, pp. 294–319.
Sutton, Jonathan (2005), "Stick To What You Know", *Nous* 39.3, pp. 359–396.
Tollhurst, William (1998), "Seemings", *American Philosophical Quarterly*, Vol. 35, No. 3, pp. 293-302.
Tucker, Chris (2010), "Why Open-Minded People Should Endorse Dogmatism", *Philosophical Perspectives* 24, pp. 529–545.
Unger, Peter (1975), *Ignorance: A Case for Skepticism*, Oxford: Oxford University Press.
Way, Jonathan & Whiting, Daniel (2017), "Perspectivism and the Argument from Guidance", *Ethical Theory and Moral Practice* 20, pp. 361–374.
Wedgwood, Ralph (2002), "The Aim of Belief", *Philosophical Perspectives* 16, pp. 267–297.
Williamson, Timothy (2000), *Knowledge and its Limits*, Oxford University Press, Oxford.

Ethical and Practical Normativity

Reasons and Oughts: Fundamentals within the Normative

OLAV GJELSVIK

Introduction

There is an ongoing debate about whether reasons, oughts, or both of them are fundamental in the normative domain. This paper is a contribution to this debate. It argues against an elegant reduction of reasons to oughts and thereby defends the fundamentality of reasons, and it can be seen as lending support to the view that both reasons and oughts are essential and fundamental.

The view that reason is the fundamental normative notion has wide support in recent philosophy. John Skorupski named his great book *The Domain of Reasons*. (Skorupski 2010.) Joseph Raz famously wrote that 'the normativity of all that is normative consists in the way it is, or provides, or is otherwise related to reasons' (Raz 1999, p. 67). Derek Parfit and Tim Scanlon also treat the notion of being a reason as the fundamental normative notion, not reducible to any other, not reducible to anything non-normative. (Parfit 2011 and 2014, Scanlon 2014.)

The view has also met opposition. The strongest opposition is from John Broome, who has argued that reasons are not a fundamental element of normativity, and that 'ought' is. In my judgement, Broome's work is argumentatively the most developed of all contributions in this area. It remains a minority view but is more tightly argued than the majority views.

My more precise aim in this paper is to provide a partial defence of the view that reasons are fundamental, without arguing that oughts are not. The point of departure will be some formulations of Derek Parfit's. While most writers state

that facts *are reasons*, Parfit as a rule does not, but maintains that facts *give or provide reasons*. What a fact provides might then be seen as something that bridges a non-normative fact and a normative fact, in turn expressed by an 'ought'.

Tim Scanlon has said explicitly that a reason is a consideration that counts in favour of something (Scanlon 2014, p. 44).[1] Without going deeply into exactly how one needs to think of 'consideration' and its ontology to make this claim fully plausible and true, I note that a consideration has the function of bridging facts and oughts, first and foremost by taking a fact as input to something that delivers support for an ought. Accordingly, the Parfitian view could be that when we say that a reason is the fact that F, i.e. the non-normative fact that F, what we mean or should mean by this is that this non-normative fact that F is essential input to a consideration that supports some ought statement. This consideration clearly needs to be conceived as having an objective existence, not mind-dependent nor response-dependent.[2]

The relation of support that the consideration realizes can in many cases be seen as something brought about by an argument from the fact in question together with other things, for instance general principles, further particular facts etc., to the ought statement in question. These things together make up the consideration that lends support to the ought statement. Note that there may be nothing like a detailed argument in the sense that a person responsive to the reason

[1] See also Scanlon 2014, p. 2, where he maintains that reasons 'are not reducible to or identifiable with non-normative truths, such as truths about the natural world of physical objects, causes and effects, nor can they be explained in terms of notions of rationality or rational agency that are not themselves claims about reasons'.

[2] Note that Parfit, especially in the second part of *On What Matters*, vol 2, where he states his metaphysical view, uses the word 'ontological' in an unusual way, a way different from my usage, namely in a sense in which the statement that 'numbers exist' does not express an ontological view. In my use of 'ontological' this statement about numbers does express an ontological view, namely that there are these (abstract) things called numbers. I agree with Parfit in that individual reasons exist in the same sense as numbers exist. While I make no distinction between existence and ontology, my view is very close to Parfit's non-reductive and non-naturalistic normativism.

needs somehow to have this argument present to her or his mind. That seems too strong. It must be enough that there is such an argument, and that it can be truly said that there is, in ways to be identified and specified, sufficient ground for thinking the person responsive to the reason. There may indeed be many ways to satisfy such a weaker requirement, as there are ways to respond that exhibit what may be considered sufficient reason responsiveness even if there is no such argument explicitly present to the mind of the person. Still reasons, on such a view, have an essential role in the normative and are fundamental by connecting the non-normative and the normative, by bridging them in considerations, considerations present to a mind or not, in some way to be manifested in thought and action.

I shall focus on Broome's arguments against the fundamentality of reasons. I first turn to his view on whether facts are reasons.

Reasons and facts

Broome writes: 'The view that reasons themselves are fundamental elements of normativity can be quickly dismissed, and I do not think many philosophers really hold it. Many reasons are natural facts. For example, the fact that apple-pips contain cyanide is a reason not to eat too many of them. Natural facts are not features of normativity at all, so they cannot be fundamental elements of it.'[3]

It is true that many philosophers seem to hold that natural facts simply are reasons. But 'are' and 'is' are words with very complex uses, as the 'is' of identity and the 'is' of predication. One may say that 'a reason is the fact that …', without thinking of this as identifying the reason and the fact ontologically. One may thus think of reasons and facts as kinds of individuals, not as to be identified with each other. One might, alternatively, think that being a reason is a property of the fact. On the general assumption that facts are countable individuals, that opens for predicating of facts that they are reasons, and count reasons the way we for instance count the red cars among the cars. This will probably open for the pos-

[3] Broome 2018, p. 1. (References are to the version on Broome's homepage.)

sibility of there being more than one kind of reason-property. This view rejects seeing the 'is' in 'is a reason' as the is of identity. I shall return to issues around this approach later in this article; it depends on a metaphysics of countable facts prior to and independent of reasons or the property of being a reason, and on seeing facts as entities we can predicate such a thing as 'being a reason' of.

We might alternatively just think that reasons are countable individuals in their own right, as we surely seem to count them when we say that there are two reason in favour and one against. This seems the natural way of interpreting Scanlon, and his invocation of 'consideration': We understand reasons simply as countable abstract considerations that make essential use of particular facts. Now, consider this example in the light of the last possibility:

The natural fact that the earth presently is heating up extremely rapidly provides a reason to believe that this heating is happening. The very same fact also provides a reason to do something to stop this heating, as it will have very dire consequences. This fact about rapid heating is not a feature of normativity, but the reason to believe that it is occurring and also the reason to do something about it are features of normativity. Therefore, these reasons, if they are countable individuals, seem ontologically distinct from the natural fact of heating. They seem to be different reasons and then they cannot both be identical with the same fact. If they do not lead to the same ought in all possible worlds, they are not the same reason. In that case the natural fact provides two reasons.

On the face of it, this little example may show several things. The first thing it may show is this: the relation of giving or providing may deliver different types of reason as the one and same natural fact may provide both a reason for belief and a reason for action. Note how the considerations that link the natural fact and the normative seem relevantly different in the two cases: The consideration in support of what you should believe is different from the consideration in support of what you should do. I shall say more about this shortly. A similar case is this: the fact that a child has fallen into a pond, provides a reason for believing that a child has fallen into the pond. It also provides a reason for jumping in to save

the child. This might indeed be what Parfit possibly had in mind by his use of 'providing'.

Evidentialism is a widely accepted epistemological thesis, and it maintains that all reasons for believing p are provided by the evidence for p. I shall accept this thesis, and make use of it, but cannot argue for it here. A closely related view is that reasons for believing p simply are the things that matter for the truth of p. There might surely be many beneficial effects of being in the state of believing p, but these effects do not provide reasons for believing p, even if they may provide a reason for being in this state of believing p. So, even if being in the state of believing that a child has fallen in hopefully has beneficial effects when a child has fallen in, these effects may be deemed irrelevant for believing that a child has fallen in. What is relevant for the belief is only evidence in support of the proposition that a child has fallen in or the things that matter for the truth of that proposition.

The natural conjecture arising out of this first point is this: There seems to be at least two quite distinct kinds of reasons provided by one fact, reasons for belief and reasons for action, as the reasons for action will always include the beneficial effects of the action in question.[4] Of course there may be further kinds, reasons for beauty, for example.

Secondly, consider again how Broome's own example of apple pips might throw light on this: Broome wrote: 'For example, the fact that apple-pips contain cyanide is a reason not to eat too many of them.' Notice that, if it happened to be true that you suffered from some rare condition that could be treated only by a minute daily intake of cyanide, and apple pips were your only source of cyanide, and eating them would help you, then the fact that the pips contain cyanide may be a reason for eating some of them.

The point is simply that a fact typically provides a reason or is a reason for doing something only *in the light of a large number of further facts*. These further facts play a role in the relevant consideration, as assumed background, presuppositions etc., without any focus upon them. The mentioned fact

[4] For a very thoughtful discussion of these matters, see Kelly 2003, and his use of a distinction between categorical reasons for belief, and hypothetical ones.

might actually be something like the tip of an iceberg of facts, but the significant bit in our context. These further facts may play a role in determining that eating the apple pips would be good or bad for you. Something similar in fact seems to characterize the presence of any reason for action. If so, the further relevant natural facts in the practical case matter in ways that seem different from the way they matter and could matter in the theoretical case of whether to believe that the apple pips contains cyanide. In that theoretical case there is only a categorical reason, namely that they do contain cyanide. This is hidden from view by the simple use of 'the reason is that they contain cyanide'.

Summing up this second point: A natural fact is a reason for doing something only relative to a number of further facts, and, as these further facts change, it can be a reason for doing something quite different.

The natural fact that apple seeds contain cyanide is a categorical reason for believing that apple seeds contain cyanide and is such a reason quite independently of all these further facts that are relevant for an action like eating the pips. And the natural fact in question is never a reason for believing that they do not contain cyanide. Reasons for belief thus work very differently from reasons for action. The differences show up in the different considerations, and correspondingly in the way a fact that is a reason for both belief and action in the actual world might fail to be a reason for action in many possible worlds where it still is a reason for belief.

This point definitely points to great caution in how to see the relation between a natural fact and reasons. There seems to be a one-many relation between a natural fact and the reasons it provides. This is hidden from view by the simple use of 'the reason is that they contain cyanide'.

Looking further at Broome's argument, it mainly draws out an implication of a view many people hold or seem to hold. But it does not provide an argument to the effect that natural facts simply are reasons or identical with reasons. If one connects identity-conditions with further features, like being given by 'sortal' predicates, then one can also say that part of the argument simply shows that '…is a reason is' is nothing like a 'sortal' predicate. But whether it is something

like a 'sortal' predicate or not is not to be settled by a majority vote. Here we need philosophical arguments.

If Parfit instead holds, as I interpret him to do, that reasons are not strictly speaking identical with facts, not properties of facts, but instead abstract objects that are provided by facts and which may serve as a bridge to ought statements, then the standing of this view is not affected by an argument that simply presupposes a different view, or by the fact that the other view in question is a majority view. Ibsen, a fellow countryman of mine, of course maintained that the majority is never right.

I am not suggesting that we should follow Ibsen in any matter. I only suggest that we see Parfit's consistent use of the terms 'providing' and 'giving' instead of the simple term 'is' as expressing caution about saying without further qualification that reasons are (identical with) facts, and as hinting at a different view that denies that we should take the statement that natural facts are reasons as identifying or revealing the very nature of reasons. Reasons are instead tentatively seen as something that in some sense takes such facts as input and deliver something with a bearing on ought statements. A reason may be provided when an input is provided but reasons *are* not the input. Again, it should be obvious from the context that if we were to call the reason a consideration, the notion of consideration would have to be fully objective, and completely independent of anything mental or features of the person engaged in considering. The objectivity of reasons' existence must not be affected by the use of 'consideration'. This is simply a different view of reasons, to be explored below, and unaffected by Broome's quick argument. 'Providing' and 'giving', in contrast to 'is', can be more easily and more clearly taken to exhibit one-many relations, relations to reasons with different properties that do not exist in the same possible worlds, to be exemplified by the cases discussed.

Reasons' and facts' place within basic metaphysical picture

Broome's main metaphysical view is that, and I quote, 'The metaphysical domain of normativity is a domain of properties and relations only. Things can have normative properties

and stand in normative relations to each other. For example, it can be the case that a particular person ought to do an act of a particular sort. But the things – the person and the act in this case – are not themselves normative.' (Broome 2018, p. 1.)

On Broome's view, there are no normative 'things', meaning individual normative entities, as reasons might be taken to be. 'Things' can, according to Broome, enter normative relations and have normative properties but no 'things' are themselves normative. I shall assess this picture, which excludes an ontology of reasons.

To make his argument from explanation, Broome introduces a parallel between the property of being a reason and the property of being magnetic. The property of being magnetic is or seems to be reducible to other and more fundamental physical properties. The property of being magnetic is in that case not fundamental. Broome goes on to say: 'But explanation goes in the opposite direction: a magnet (though not the property of being magnetic, note) explains the tendency of objects containing iron to move towards it'. (Ibid.)

I shall not here challenge Broome's suggested view on explanation, even if I do not share it. I shall focus elsewhere: Broome makes use of a basic parallel between being a magnet and being a natural fact. The parallel lies in the way a magnet (but not the property of being a magnet) can enter explanatory relations, and with how a natural fact can enter explanatory relations. The property of being a reason, a property of the natural fact, would have a role parallel to that of the property of being a magnet. This requires that the property of being a reason is no more fundamental than the property of being magnetic.

The question is then whether the presence of such a natural fact stands in some sufficiently analogous way to the presence of a magnet, as something we should recognize as an ingredient in our ontology just as magnets are.

Consider the question of whether we have a theoretical need to make a commitment to an ontology of natural facts on general or independent grounds, a commitment able to give facts an ontological standing sufficiently parallel to that of magnets, making them able to enter explanatory relations in a sufficiently parallel way. Both magnets and natural facts would in that case be 'things' we commit to quite inde-

pendently of whether the corresponding property of interest is reducible to other properties.

Let us take for granted that we want to commit ontologically to what important explanations commit us to, and that we do not challenge Broome's somewhat special view on explanation. Then, if there is no other ground for commitment to facts beyond their role in providing reasons that enter explanations of oughts, then this seems to mark a possible difference to the magnet case. We might not have any independent general grounds for committing to an ontology of facts – whether we have such grounds is, at least so far, an open question. Maybe we can make do with true propositions (no ontology of facts, just an ontology of propositions) in all contexts besides this present one where there seems to be a real explanatory role for facts/reasons (according to Broome). If it is facts' having the property of being reasons or providing reasons and in virtue of that having an essential role in explanations of oughts that introduces a need or a ground for the ontological commitment to facts in the first place, a commitment that is required by the parallel to magnets, then we seem to have a difference to the magnet case.

In the magnet case, one can rest content with an ontology of magnets and a still accept a full reduction of the property of being a magnet. The ontology of the things that are magnets has further grounds, as the magnets we speak of are physical objects in space and time with many further physical properties. The possibility we are facing is this: Facts having the property of being a reason seems possibly to make up the entire ground for the ontological commitment to individual facts and thereby to reasons. In contrast, having the property of being a magnet does not in itself provide the full or only ground for the ontological commitment to the objects in space and time with many physical properties that also have the possibly reducible property of being magnetic.

Conclusion: If we should accept an ontology of facts only out of them being reasons or providing reasons, then there may be an asymmetry with the magnet case that Broome has to deal with. This asymmetry relates to the issue of whether

there is a need for an ontology of individual facts.[5] The possibility of this asymmetry and its significance is not discussed by Broome. It makes it hard for him just to appeal to the parallel, if an ontology of reasons (and also facts) only arises out of their role in these explanations of oughts. The issues raise many questions about how to think about metaphysical commitment, with strict criteria we should not commit to the facts needed for seeing 'being a reason' as a property of facts, with much looser criteria there may be room for both facts and reasons, as different kinds of abstract individuals.

Reduction: Can talk of individual reasons be reduced to talk with the mass-term 'reason'?

With this point in mind, I turn directly to another crucial issue, and Broome's intermediate claim that 'reason' can be seen as a mass-term. Broome's overall strategy is not entirely dependent on making this claim, but it is essential for one line of reasoning he advocates. I turn to this line of reasoning, and the relationship between object-terms, mass-terms and the predicate of being reasoned-for.

As noted above, we sometimes clearly speak of individual reasons, for instance that there are three reasons for something, and one against. We sometimes make the claim that this is the same reason as we have encountered before, and sometimes that this is a new and different reason even if it has similarities with a previous reason at first glance. At other occasions, we seem to speak of the stuff of reason, i.e. we use the mass term 'reason', saying that there is much reason for you to believe that the earth is heating up. Broome's more fundamental metaphysical view is that we are better off by constructing a way of speaking about reason and individual reasons that avoids altogether both reason stuff and individual reasons. That view, he claims, can be reached by a reduction in two steps. The first step is a reduction from talk of individual reasons to talk of the reason stuff, the second step takes us from talk of reason stuff to a new artificial predicate.

---

[5] One of the best cases for that is in my judgement given in Neale 2001. For an argument to the effect that Neale's arguments are far from sufficient, see Gjelsvik 2006.

This new artificial predicate, 'is reasoned', is explained in this way, by this example:

> 'The lights coming on at dusk is reasoned' means just the same as 'There is reason for the lights to come on at dusk'.

Note that here it is the mass-term 'reason' that does the explaining in relation to the new predicate 'is reasoned'. And the use of the mass-term has the same extension as the statement that the weight of all relevant reasons is on the positive side for the light to come on at dusk.

There is also another new artificial predicate in this neighbourhood, namely the predicate 'is oughted':

> 'The lights coming on at dusk is oughted' means just the same as 'The lights ought to come on at dusk'.

With the first predicate available, Broome can say what he wants to say without committing himself to reasons or even to the stuff referred to by the mass-term reason. He can make do with the artificial predicate 'is reasoned'. His further points hold as long as he, with both these artificial predicates available, can reduce the artificial 'is reasoned' predicate to the new 'is oughted' predicate.

Here are some difficulties:

These artificial predicates are not predicates of English or of any other language. In order to explain their meaning, we have, at least in one of the cases, the 'is reasoned' case, to use and employ terms that seem to commit us ontologically to individual reasons or to reason stuff.

There seems to be ontological simplification and gain in the artificially constructed predicates. But whatever ontological gain there is in the artificial construction may seem lost if we have to refer to the stuff reason or to individual reasons in order to explain the meaning and use of the artificial predicate.

Broome wants to first make do with the mass term reason, and then reduce the use of that to the artificial predicate. In order to do that he clearly needs to succeed in the first step, the reduction of individual reason talk in favour of the mass term.

Explaining individual reasons from an employment of the mass term 'reason'

With this in mind, let us turn to this crucial and necessary step, namely to the relationship between talk of individual reasons and talk of reason stuff, between there being a reason and there being reason. It is true that important philosophers have tried to explain what *a reason* is by providing an explanation out of the mass term 'reason'. I turn to this.

Here is the main example.

> Thomas Nagel's definition: Every reason is a predicate *R* such that for all persons *P* and events *A*, if *R* is true of *A*, then *P* has prima facie reason to promote *A* (Nagel 1970, pp. 47-48.)

There are several issues here. A reason is not a predicate, but let us push that objection aside, as a simple mistake that can be corrected for. The further idea is, or seems to be, that the count noun reason can be explained from the mass term 'reason'. This is an idea Broome endorses for the purposes of the reductive enterprise. If this idea holds up, the tiniest bit of the stuff denoted by the mass term 'reason' is enough for there to be a reason. This definition of a reason thus gives no clue at all as to how many reasons there are for or against something, so there is not here available anything like a definition of the count noun, or individuation of the things denoted by the count noun. Since this is so, there simply is no reduction of individual reason talk to mass term talk, and it may surely be more natural to reduce the mass term talk to the talk of individual reasons. But the proposed definition can possibly be seen as connecting there being reason for A and support for promoting A.

Broome's strategy in this part, and I repeat that this is not his only strategy towards his main goal, starts from a way of taking the mass term as explanatorily fruitful in relation to explaining the meaning of the count noun, and then provides a reduction of the mass term to a predicate that applies to what reason supports. Thus, some things get to be 'reasoned' (when there is reason to promote them), which again is a predicate explained out of the employment of the mass term 'reason'.

Parfit and Scanlon stick with the count noun reason, and do not define the count noun out of some employment of the mass term.

Is 'reason' a quasi or fake mass term?

Here we encounter an issue not discussed by Broome: The possibility of fake mass terms or quasi mass terms. Well discussed examples are things like 'furniture' or 'cutlery'.[6] These terms look like mass terms but are not really mass terms. There is, in these cases, no ontological commitment to the stuff, only to individual pieces of furniture or to spoons, forks and knives. Reason also has a mass term use, as well as a count noun use. The question is then whether the mass term 'reason' is a fake mass term. In that case there may be no ontological commitment to the stuff, and the use of the mass term cannot be taken at face value when doing ontology. Note that the term 'reason' is very interesting because it comes in both a mass term use and a count noun use. This is unusual.

Fake or quasi mass terms are thus cases where the count noun really is the primary term when it comes to ontological issues, even if it may look otherwise at first glance. Consider in some detail the case of the fake mass term 'cutlery'. We clearly think of the individual pieces of cutlery as the only ontological commitment at the metaphysical level, the mass term 'cutlery' is thus derivative, metaphysically speaking. The individual pieces of cutlery can be counted, and if there is no individual piece of cutlery, then there is no cutlery. This is so even if there is stuff or parts one can make cutlery out of.

The phenomenon of quasi or fake mass terms is very widespread in my own language, where it often just represents a stylistic variation, without metaphysical significance. In this stylistic variation, we get something like a double use, both as count noun and as mass term. When we say that there is much blueberry this year, the ontological commitment is that there are very many of them this year. If there for example were only one but an absolutely gigantic 'blueberry', we

[6] For an introduction to these issues, see for example the entry in *Stanford Encyclopedia of Philosophy* on 'The Metaphysics of Mass Expressions' by Mark Steen (2016).

could not easily use the mass term ('much blueberry') to correctly describe the situation even if there is much blueberry stuff. This might of course be because we have a hard time thinking of the gigantic berry as a blueberry, but the most likely ground for our hesitance is that 'much' here simply in some way stands for 'many'. And if there only existed blueberry mash or blueberry juice scattered all over an area, we could not say that there is much blueberry in that area, even if there is much 'blueberry stuff'. The same applies if we say, which we can do in my language, that much men are out tonight. (In English, 'much people' seems possible.) If only one absolutely gigantic man, a man the size of many normal people, is out, and everybody else is in, this will not easily make the statement that much men are out true. Even if we think that the giant is a man, we might still hold back, and the reason is again that 'much' here just is a way of speaking about many. If there were lots of body parts and blood scattered around, but no people, we could not say that much people were out. The case is thus reasonably parallel to the case of blueberries. There is any number of such cases.

Now, the issue is whether the statement that there is reason in favour of something S is a statement made true by there being at least one consideration in favour of S, or whether it rather refers to some stuff that may be there independently of all normal individually countable reasons or considerations. It seems to me that the latter is false, and that it follows that an ontological commitment to reason stuff has little or no attraction. However strong or minute a reason, the stuff is not there independently of individually identifiable reasons. We could therefore, at least in theory, accept Broome's view about the abstract possibility of reducing 'there is reason' to 'is reasoned'. However, as long as we reject the metaphysical reduction of talk of individual countable reasons to the talk of reason stuff, that will block the overall reduction Broome needs. It also seems that in all normative explanations, explanations of the obtaining of ought, we start from individual reasons, given or provided by individual facts.

It is of course true that we often weigh reasons, and that there is some limited analogy between reasons and physical forces. We say for example that this is an important reason,

but that that is not an important reason, and that three unimportant reasons in favour of something can easily be outweighed by one important reason against. But unless there are considerations that give some further content as to how we can get out of talking about individual reasons, considerations that can throw independent light on what we mean by speaking of reason stuff as denoted by the mass term reason, and we do not really have any idea of what such reason stuff is or what such talk may amount to metaphysically beyond talk of individual reasons, then there is a challenge. We only seem to know what reason stuff is or may be from our knowledge of individual reasons that for sure can enter complex relations with each other and be weighed against each other. Results of weighing we might express by saying that reason counts in favour of, there is more reason for than against etc. But that seems to be all, and an argument seems required for this to be construed metaphysically.

Now if we, as I recommend, also recognize at least two kinds of reasons, epistemic reasons and practical reasons, where the kinds in questions are characterized by the fact that there is no meaningful way to weigh them against each other, and thus no way we conceive of these kinds of reason just being of the same reason stuff, then the intermediate step of metaphysical reduction to one sort of stuff seems very hard to defend. And it has to be defended to reach Broome's full reduction to ought via the general reduction of 'is reasoned' to 'is oughted'. Broome could of course claim that also he sees 'reason' as a fake mass term, since his view is also a reduction of the mass term. But if he cannot make good on the first reduction, he cannot make good on the final reduction, and then he is not entitled to such a claim.

'Is reasoned' and 'is oughted'

Let us now turn specifically to the predicate 'is reasoned' and make similar points in relation to that. Caroline needs a loan, and this fact provides Caroline with a reason to go to the bank. Her going to the bank is then 'reasoned', by Broome's definition. Note that this predicate of 'being reasoned' is true of 'Caroline going to the bank'; for it to be true Caroline must have a reason for going to the bank. The reason itself is pro-

vided by the fact that Caroline needs a loan, or it is the fact that she needs a loan.

As we have seen, Broome's strategy is to get rid of both reasons and reason stuff in favour of the predicate being 'is reasoned'. If we from the related fact that Caroline ought to go to the bank, i.e. 'Caroline going to the bank' being 'oughted', can provide the truth conditions also for 'Caroline going to the bank' being 'reasoned', then a reduction of the latter predicate can be carried out. Broome also maintains that we cannot make do without the predicate ought, so an eventual symmetry between 'reasoned' and 'oughted' in the explanation of truth conditions would make no real difference for what we should treat as fundamental out of the two. From this it seems to follow that fundamental normative predicate has to be 'is oughted' and ought.

The first question is then how we can reach a determination of the meaning of 'is reasoned' so that we can really assess the claim that it means the same as 'is oughted'. It now seems that such a determination has to start from an employment of the countable noun 'reason', and from that establish or reach an understanding of the use of the mass term 'reason'. This is necessary in order to make it possible to discuss whether the truth conditions of 'is reasoned' and 'is oughted' really are the same. Now, if we really need, as I claim, to start from individual countable reasons in order to determine and explain the meaning and reference of the mass term 'reason' and from that move on to a determination of the meaning of the artificial predicate 'is reasoned', and no ontological reduction of individual reasons to the ontology of the mass term is defensible, then the theoretically possible move from the mass term and its ontology to the predicate 'is reasoned', which only seems to get rid of the ontology of the mass term, the stuff 'reason', will not help us all that much. This is so as long as the vocabulary and talk of individual countable reasons is necessary to provide the meaning there is or might be to both the reason stuff talk and the further predicate that is constructed out of the stuff talk.

If this is right, then the arguments about the identity of the truth-conditions of the predicate 'being "reasoned"' and the predicate 'being "oughted"', opens for looking at things differently. It might even invite a possible reduction of 'is

oughted' to individual reasons. However, if there are contexts in which we in no way can do without the predicate 'is oughted' or simply ought, as Broome claims, then the overall picture seems to leave us with two fundamental notions: reason and ought. The necessary commitment to the reason bit seems to commit us to individual, countable reasons, and the main result so far is that there is no way easily available for getting out of that commitment. If individual reasons have facts as input, and are considerations that also make use of principles with oughts in them, both notions are fundamental.

Concluding overview

The strategy of reducing talk of reasons to talk of oughts thus faces real difficulties. Parfit or Scanlon have not pointed out these difficulties, nor have they defended their views against the tightly argued criticism by Broome. Some difficulties with Broome's strategy now stand out. I shall put some of them together in a final picture.

Consider this again

- The fact that apple pips contain cyanide is a reason to believe that apple pips contain cyanide.

- The fact that apple pips contain cyanide is a reason not to eat them.

Broome aims to reduce 'being a reason' to 'being oughted'. He does that by moving from the property of being a reason to the predicate 'is reasoned', and from that to the predicate 'is oughted'.

He also thinks that we need an ontology of facts to account for how reasons explain the obtaining of oughts, but no further ontology of individual reasons. In this connection, he also argues that we can ascribe the property of being a reason to facts, and from that move on to the predicate 'being reasoned', and from there on to the predicate 'being oughted'.

One counterargument claims that we may not get rid of individual reasons. It starts from observing that it is not clear that we should have an independent ontology of facts if they do no explanatory work beyond what true thoughts or propositions do. Consider this: If facts as such do no explanatory

work in semantics or in our account of representation, and they are admitted into in our ontology only because individual reasons are needed for work in explaining the obtaining of normative relations or oughts, then what? In that case we do not get rid of individual reasons in our ontology if the facts they are supposed to be properties of are admitted into our ontology only in so far that they are reasons. Reasons sneak in the back door if facts only exist in so far that they are reasons.

Secondly it has been argued that the route through employment of the mass term reason, used to define a reason as Thomas Nagel did, is blocked because reason is not a real mass term, it is a fake mass term. The ontological issues are then being referred back to individual reasons.

Putting this line of argument aside there is another argument, a third type of consideration. This is a methodological objection concerning what implications can be drawn from the view many people hold to the effect that facts are reasons. If one simply holds a very different view, to the effect that facts are inputs to reasons, that reasons take this input and deliver something that contributes to the truth of ought statements, that we identify reasons by their significant input, and that reasons can be understood as considerations that actually or potentially are present to the mind of the agent, and that these consideration in some paradigmatic cases can be laid out as an argument or as a representation of the weighing that some way or other is internalized and thus takes you to a judgement about what you ought to do, then this different view is not effected by Broome's quick argument.

This alternative view thinks very differently about what reasons are, which I also suggest may be behind Parfit's use of language.

Then take again the fact that apple pips contain cyanide. This fact seems to provide two reasons, one for belief and one for action. These reasons are not obviously commensurable and cannot be reduced to the same kind of reason stuff. They do not even exist in the same possible worlds. In all worlds where the fact in question exists it is a reason for believing that apple pips contain cyanide. In many worlds where this fact exists, it is a reason against eating the pips, but in some

other worlds it is a reason for eating the pips. Reasons for belief and reasons for action thus come apart across possible worlds, as there being reasons for action always depends on a large number of further facts that reasons for belief do not depend upon.

A natural fact may therefore provide different individual reasons, reasons with different properties which may exist in different worlds. This may be because individual reasons are not facts but take facts as input in their own special way in different settings. It is in that case natural to introduce the relation of giving or providing reasons. If reasons are considerations that bridge facts and the obtaining of oughts or contribute to such bridging, then such considerations and such bridges will also be of at least two basic sorts. This is because the considerations, or the potential 'argument' that provides the justification that is at the bottom of the 'explaining' in question, is different in the epistemic case and the practical case. The accurate term is then that facts provide reasons, where providing can be one-many. Secondly, within the considerations that support oughts, there may surely be employments of other oughts, and then both notions, reason and ought, are fundamental.

I will end with a further point I have not discussed so far, not to make things too complex. But it needs mentioning. When Broome moves from talk of reasons and reason to the predicate 'is reasoned', which is claimed to have the same truth-conditions or be true in the same circumstances as 'is oughted', there also seems to be a risk of reducing reasons to their overall net effect. But that would also seem to be wrong. We acknowledge the existence of a specific gravitational force between two objects even if there are many other forces around that also influence the actual behaviour of these objects. Furthermore, we do seem to treat individual reasons in a way similar to how we in this case treat forces. The reduction of forces to a final net force is highly relevant for predictions but does not settle the ontology of forces. Even if we only granted the existence of some sort of mass 'force', in all concrete situations simply the net resultant force, it is not clear why we should want to get rid of the further ontology of kinds of force and replace it with some predicate 'is forced'. The reduction of forces should only result from a uni-

fication of them within physics. That may of course be forthcoming but is an entirely different matter.

All in all, this shows the suggested general reduction of 'being a reason' to 'ought' faces real problems. As far as we have come, we should acknowledge that there is a good case for seeing reasons as fundamental within normativity; ontologically speaking they are *sui generis* and play a fundamental role within the normative. That does not by itself make them the only fundamental notion within normativity.

Acknowledgement

I am very grateful for comments on this work from the audience at the IIP meeting in Helsinki, and also to an audience at the University of Oslo. I thank Tim Williamson for comments on the claim about asymmetry between the magnet case and the reason case. I am especially grateful to John Broome for a long and generous discussion about these issues at a coffee shop in Oxford. Finally, I acknowledge help from valuable comments from two anonymous referees.

University of Oslo

References

Broome, John (2013), *Rationality Through Reasoning*, Wiley-Blackwell.

Broome, John (2018), "Reason fundamentalism and what is wrong with it". In *The Oxford Handbook of Reasons and Normativity*, edited by Daniel Star, Oxford University Press, pp. 297–318. Also on John Broome's website, University of Oxford.

Broome, John (Forthcoming), "Giving reasons and given reasons". Forthcoming in a Festschrift for Derek Parfit, edited by Tim Campbell, Jeff McMahan and Ketan Ramakrishnan, Oxford University Press. Also on John Broome's website, University of Oxford.

Gjelsvik, Olav (2006), "Facing Facts and Motivations", *Protosociology*, Vol 23. *Facts, Slingshots and Anti-Representationalism. On Stephen Neale's Facing Facts*, edited by Gerhard Preyer and Georg Peter, pp. 259-271.

Kelly, Thomas (2003), "Epistemic Rationality as Instrumental Rationality: A Critique." *Philosophy and Phenomenological Research* 66 (3), pp. 612-640.

Nagel, Thomas (1970), *The Possibility of Altruism*, Oxford University Press, Oxford.

Neale, Stephen (2001), *Facing Facts*, Oxford University Press, Oxford.
Parfit, Derek (2011), *On What Matters, vol 1*, Oxford University Press, Oxford.
Parfit, Derek (2014), *On What Matters, Vol 2*, Oxford University Press, Oxford.
Raz, Joseph (1999), *Engaging Reason: On the Theory of Value and Action.* Oxford University Press, Oxford.
Scanlon, T. M. (1998), *What We Owe to Each Other*, Harvard University Press, Cambridge.
Scanlon, T. M. (2014), *Being Realistic about Reasons*, Oxford University Press, Oxford.
Skorupski, John (2010), *The Domain of Reasons*, Oxford University Press, Oxford.
Steen, Mark (2016), "The Metaphysics of Mass Expressions", *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/metaphysics-massexpress/>.

"A Bewildering Conceptual Jungle": von Wright's Analysis of the Ambiguity of the Term "Norm"

SARA HEINÄMAA

In contemporary philosophy, the terms "norm" and "normative" are used in numerous thematic and argumentative contexts. Both terms operate in several different meanings and have several different connotations, such as standard, measure, guide, goal, model, optimum, limit, condition. Even if the basic ideas of rule-following and correctness are implied by most usages of these terms, the types of rules and the types of correctness at issue vary greatly depending on the debate and its main topic. Further, the large coverage of the topics – from the general structures of human conduct and action to the most demanding aspects of our epistemic, aesthetic and moral lives – suggests that the variance of meanings may be too broad for theoretical and systemic purposes. This worry is expressed pertinently by Stephen Finlay in his recent essay "Defining Normativity":

> In recent jargon, *metanormative* theory explores fundamental questions crosscutting ethics, political and legal philosophy, aesthetics, epistemology, and more. It is described as the study of *normativity*, suggesting there is something, called 'normativity', that is the common object of the competing theories of the philosophers working in this field. The literature on 'normativity' has in a short time become overwhelmingly huge. So, a curious layperson might reasonably ask, 'So, what is this 'normativity' then?' This innocent little question might already be interrogation enough to make philosophers squirm and sweat, because it is hard to find any definition that every metanormative theorist can agree on [...]. At least one leading practitioner, Derek Parfit has recently gone so far as to claim that many philosophers who

> appear to disagree with him about the nature of normativity must be using their terms with different meanings, and talking about something else entirely. [...] This paper [so Finlay's] finds, in partial agreement with Parfit, that philosophical discussion about 'normativity' is plagued by systematic ambiguities contributing to significant confusion, as there are many things that 'normativity' can reasonably be taken to be (Finlay 2018; cf. Parfit 2011).

If Finlay's analysis holds, then we need to make distinctions and add specifications to the terms "norm" and "normativity" in different contexts of argumentation. For this purpose, I turn to a philosophical discussion which may seem remote and already bypassed or surpassed. In order to make sense of the multitude of usages given to the terms "norm" and "normative" in contemporary philosophical debates, I want to revisit Georg Henrik von Wright's Gifford Lectures from 1959–1960, published as *Norm and Action: A Logical Enquiry* (1963). Even if the source is relatively old, its distinctions are still helpful, I think, in organizing the conceptual field of norms and normativity. Some of von Wright's distinctions are widely adopted in certain argumentative contexts, for example in jurisprudential and social-ontological discussions (e.g. Lorini & Zelaniec 2018; Passerini & Di Lucia 2017), while others are largely forgotten or ignored. To widen the perspective, I want to present the analysis as a whole and, at the end, draw some parallels to phenomenological philosophy.[1]

1. "The jungle"

According to von Wright, the term "norm" is used in six different senses in ordinary, scientific and philosophical discussions. By "norm" we can mean either (i) an enabling rule, (ii) a prescription, (iii) a directive, (iv) a custom, (v) a moral principle, or (vi) an ideal principle. Von Wright argues that all these six categories can be said to involve "norms" but in crucially different senses. So, what we have is not just six different specifications of one general concept of norm but a

[1] I discuss phenomenological approaches to normativity more comprehensively in Heinämaa (2019).

peculiar type of ambiguity.[2] Thus, the field of norms is not just superficially obscure but is a "bewildering conceptual jungle" that needs clarification in order to allow for proper philosophical agreements and disagreements (von Wright 1963a, p. ix).[3]

More precisely, von Wright argues that the term "norm" is not genuinely ambiguous as is, for example, the English term "bank", which means both a financial institution and an edge of a river, or the term "duck", which means a waterbird, a type of fabric, and a score in cricket. Unlike genuinely ambiguous terms, von Wright argues, the term "norm" has several different meanings, yet meanings that are related to one another by complex logical-semantical connections. This means that even if a philosophical explication of the terminology of norms is urgently needed, the attempt to create a general theory covering the whole field in all dimensions would be futile (cf. von Wright 1963a, p. 10; Finlay 2018).

---

[2] The situation is thus logically similar to that of the term "sign" according to Husserl's early analysis in his *Logical Investigations*. In *Phenomenology: Between Essentialism and Transcendental Philosophy* (1997), Jitendra Nath Mohanty emphasizes that for Husserl, expressions (*Ausdruck*) and indications (*Anzeige*) are not two different genera of signs (*Zeichen*) but are signs in two different but logically related senses (Mohanty 1997, p. 70; cf. Husserl 1984, 30ff.). In a similar manner, enabling rules, prescriptions and goals are not three different genera of norms but norms in three different but logically related senses.

[3] Von Wright uses the metaphor of *conceptual jungle* in the preface to *Norm and Action* to characterize the logical semantics of certain norms that he calls "directives" or "technical norms" in line with Kant. He writes: "The main topic of study in this book is *prescriptions*. Originally, I had planned to include in it also a fuller treatment of that which I call *technical norms* about means to end, and the closely related topic of *practical inference* (necessity). But I have come to realize that *this is an even more extensive and bewildering conceptual jungle* than the topic of prescriptions" (von Wright 1963a, p. ix; last italics added). As the comparative makes clear, the point is not merely to suggest that the field of directives or technical norms involves ambiguities and indeterminate boundaries but that this particular area of normativity is even more complicated and confused than the other sections. Thus, the field of normativity as a whole can be said to be a conceptual jungle and the area of directives can be characterized as one of its most densely twisted thickets.

If this also holds for contemporary philosophical debates and theorization, as Finlay's analysis suggests, then we would need more than one term to discuss the involuntary tendencies of human behavior, on the one hand, and the correctness of human action and interaction (moral or epistemic), on the other. And further, more terms would be needed for the philosophical discussion of the normative structures of beliefs, cognitions, emotions, aspirations, virtues and vocations.

Von Wright's manner of philosophizing about the different senses of normativity and goodness in *Norm and Action* and *The Varieties of Goodness* can be characterized as, broadly speaking, Wittgensteinian.[4] New research on his manuscripts and essays in the 1950s has made clear the indebtedness of these two works to Wittgenstein's late philosophy (e.g. Jakola 2020; Venturinha 2020).

The methodological idea that guides von Wright's enterprise is Wittgenstein's view according to which the task of a critical investigator is to bring theoretically invested words back to their "home places", that is, to their concrete and original contexts of usage (Wittgenstein [1953] 1997, §116). In Wittgenstein's account, many theoretical-philosophical problems issue from terminological confusions, not from states of affairs, worldly or transcendental. Words have lost contact with their proper contexts of usage and actual senses and float free, so to speak, between various disputes and debates. This terminological looseness creates quasi-problems, and extensive discursive exchanges for the solution of such problems.[5] The first task of the critical investigator is to illuminate

[4] *The Varieties of Goodness* illuminates the debt that von Wright's philosophical reflections on goodness and normativity owe to Wittgenstein's later investigations. In the preface, von Wright states that he does not agree with Wittgenstein's characterization of goodness as a family-resemblance concept but at the same time emphasizes that he finds the critical tenor of Wittgenstein's inquiries crucial to all philosophy suffering from universalizing aspirations: "[T]he insight into the family-character of a concept may make us give up an attempt to hunt […] for a common feature of all things falling under this concept which would explain to us why these things are classified together" (von Wright 1963b, pp. 15–16).

[5] In *Culture and Value*, Wittgenstein writes: "I cannot found a school because I do not really want to be imitated. Not at any rate by those who publish articles in journals" (Wittgenstein [1977] 1980, p. 61).

and clarify the grounds of such formations.[6] Unlike Wittgenstein, von Wright does not give up the goals of theory construction but he sees the importance of constraining the unifying tendencies of theorization by critical philosophical inquiries (e.g. von Wright 1963b, pp. 15–16; cf. note 12 on page 14).

Inspired by this Wittgensteinian principle, I want to revisit von Wright's *Norm and Action* and *The Varieties of Goodness* and offer, on their basis, a set of philosophical-semantic distinctions that together illuminate what all can be meant by the terms "norm" and "normativity".[7] I have chosen to pro-

---

[6] In general, Wittgenstein's critical investigations work against misguided aspirations for philosophical theory-building (e.g., [1953] 1997, §109). In §118, he asks: "Where does our investigation get its importance from, since it seems only to destroy everything interesting, that is, all that is great and important? (As it were all the buildings, leaving behind only bits of stone and rubble)." His answer makes clear that the constructions that his critique dismantles have little if any permanent worth: "What we are destroying is nothing but formations of wind [*Luftgebäude*] and we are clearing up the ground of language on which they stand" ([1953] 1997, §118, translation modified, cf. §111).

The original German term that Wittgenstein uses here is "Luftgebäude" which means a mechanical formation, made by wind. Thus, his argument is that philosophical theorization largely consists of linguistic confusions and is not supported by sufficiently clear insight and free deliberation. In §133, we read: "The real discovery is the one that makes me capable of stopping doing philosophy when I want to. – The one that gives philosophy peace, so that it is no longer tormented by questions which bring *itself* in question." By this Wittgenstein does not suggest that there would not be proper philosophical problems. On the contrary, he argues that one particular species of philosophical sickness is the loss of all problems: "Some philosophers (or whatever you like to call them) suffer from what can be called 'loss of problems.' Then everything seems quite simple to them, no deep problems seem to exist anymore, the world becomes broad and flat and loses all depth; and what they write seems immeasurably shallow and trivial" (Wittgenstein [1967] 1970, §456).

[7] In the Preface to *Norm and Action*, von Wright explains that the problems encountered by his early deontic logic of obligations and permissions (1951) had led him to new reflections in the hope of establishing a more solid approach with tools more suitable for the analysis of action and its dynamism. So, despite the flourishing of deontic logic, he saw the need for a new beginning. However, the scope of this new enterprise was not to

ceed in this manner, since I find von Wright's critical discussions of the ambiguities of the term "norm" helpful for the clarification and organization of contemporary philosophy of normativity. Thus, I will not follow his analyses into specific norm types but want to explicate his general account of the semantics of normativity which prepares the ground for the analyses of *Norm and Action*.

Von Wright first distinguishes between three *principal senses* in which the term "norm" can be used. "Norm" can mean (i) an enabling rule (also: constitutive rule), (ii) a prescription, and (iii) a directive (also: technical norm). What is meant by

---

construe a general theory of norms or normativity but to focus on the logic of prescriptions and practical inference. In this context, he characterizes the difficulties of his earlier approach as fundamental: "The building of Deontic Logic has thus turned out to be a much more radical departure from existing logical theory than I at first realized. The more I have become aware of the complications connected with the subject, the more have I been compelled to narrow my claims to be able to treat it in a systematic and thorough way. What is here accomplished, if anything, covers only a small part of the ground which has to be cleared before Deontic Logic stands on a firm footing" (von Wright 1963b, p. ix). The first chapter of *Norm and Action* substantiates this decision by reference to the semantic complexities of the field of normativity: "Since the field of meaning of 'norm' is not only heterogenous but also has vague boundaries it would probably be futile to try to create a General Theory of Norms covering the whole field. The theory of norms must be somehow restricted in its scope" (von Wright 1963b, p. 1).

By distinguishing between six (to eight) components of prescriptions, von Wright extracts a threefold "norm-kernel" that includes the character, the content and the condition of application of the prescription (von Wright 1963a, p. 70; cf. 1963b, p. 157). He then hypothesizes that also other norms include identical or similar kernels and, on the basis of this, proceeds to put forward his new deontic logic as a logic of norm-kernels: "The norm-kernel is a logical structure which prescriptions have in common with other types of norms. There may however exist specific differences between the kernels of norms of different types. Here we are directly concerned with the kernels of prescriptions only. […] The formal theory of norms or Deontic Logic which we are going to develop in later chapters of this work is essentially a *theory of norm-kernels*. Since the kernels are the common ingredients of all or nearly all types of norms this formal theory may with some caution be regarded as a 'basic logic' of norms in general" (von Wright 1963b, p. 70, italics added).

"norm" in all these main cases is some kind of direction or regulation that guides human action or behavior (or human ways of relating) in one way or another. However, the sense of directing and regulating is essentially different in these three cases.

Further, von Wright then argues that by taking into account these three primary senses of the term "norm", we can also talk about (iv) customs, (v) moral principles, and (vi) ideal principles as norms. However, when we thus extend the usage of the term "norm", we must keep in mind that some of the main ingredients of normativity, identified in the three principal cases, may be missing and that correspondingly some new aspects may become prevalent.

In the following, I will first examine von Wright's three principal usages of the term "norm" and then clarify also the three additional usages.

2. Enabling rules

The first concept of norm that von Wright distinguishes is that of an *enabling rule* (also: constitutive rule). This is a norm that establishes or institutes an activity and/or practice. What is essential to enabling or constitutive rules is that they determine the steps or moves of the activity at issue and thereby define the activity itself (von Wright 1963a, 6–7). A paradigmatic example here is the rule of a game, for example, chess, Klondike or basketball. All games are characterized by normative concepts such as "correct", "permitted", "prohibited", and "obligatory". However, a failure to follow the rules of basketball or those of chess is very different from the failure to obey a military command or the failure to follow the directives of use attached to a new medicine. Moreover, all these types of failures differ from the epistemic and perceptual "failures" of not capturing all relevant aspects of an object as well as from the morally relevant cases in which one ignores an appeal or plea.

Two other cases of enabling rules are epistemologically more interesting and more central than the rules of games. These are the rules of grammars and the rules of logical and

mathematical calculi.[8] Both resemble rules of games in delimiting forms of activity *in toto*.

Von Wright emphasizes that failures to act according to the enabling norms of games, languages, and logical and mathematical calculi are characterized by a peculiar kind of duality of two alternative interpretations. Breaking the rules of a game can be understood in two alternative ways: we can either say that the person makes a mistake in playing and plays incorrectly or else we can judge that she is not playing at all, not taking part in the relevant activity at all (even if she seems to make the characteristic moves or steps):

> Of a person who does not speak according to the rules of grammar we say either that [s]he speaks incorrectly or that [s]he does not speak *that language*. The grounds of saying the one or the other are very much the same as the grounds for saying of a person either that [s]he plays the game incorrectly or does not play *it* [this game] at all (von Wright 1963a, p. 11).

There is also an important difference between the enabling or constitutive norms of natural grammars and those of mathematical and logical calculi: whereas norms of grammars have long and thick histories, the histories of calculi are "poor" (von Wright 1963a, p. 7). This essential historicity of linguistic rules is bound to their dynamic character; the enabling norms of grammars are in a constant process of slow "natural" change. This does not entail that changes in grammars would be decided in negotiations between language users or established by commands of linguistic authorities. Notwithstanding their deeply social and cultural character, norms of languages are not coined, instituted, or established by any particular people. Rather, they come about in dynamic language use as such. This means that human languages have an "organic character" (von Wright 1963a, p. 7).

3. Prescriptions

The second category of norms distinguished by von Wright are *prescriptions*. Paradigmatic cases include national and in-

[8] For Husserl's account of the norms of logic and their different dimensions, see, e.g., Hartimo 2018; 2020.

ternational laws, military commands, traffic rules and parental directions.[9] These are norms issued or given by an agent and addressed to and directed at other agents. Von Wright calls "norm-authority" the prescription-issuing party and "norm-subject(s)" the party to whom the norm-authority gives the norms.

The norm-authority is a willing agent of some sort, individual or collective. The king, the people and God are well-known examples of such authorities. Prescriptive norms can be said to have their "source" in and "flow" from the will of the norm-authority; the authority can be said to "want" the subject(s) to adopt a certain conduct or form of action; and the giving of the norm can be said to "manifest" the authority's will. Thus, the framework of prescriptions is, in von Wright's analysis, fundamentally volitional or conative.

In addition to this authority-subjects relation, prescriptions also have two other distinguishing features, in von Wright's analysis. First, they are promulgated or promoted on the part of the law-authority by symbols and marks, for example by written documents or by repeated forms of speech. This is a crucial aspect of prescriptive norms: the authority needs to promulgate and distribute her prescriptions in order to make her will known to her subjects. Second, prescriptions always come with sanctions or threat of punishment of some sort: "In order to make its will *effective* the authority attaches a sanction or threat of punishment to the norm" (von Wright 1963a, p. 7; italics added; cf. pp. 125–126).

Prescriptions differ from the enabling norms of language and logical calculi in all four respects: whereas a prescription necessarily involves a norm-authority and norm-subject(s) as well as systems of distribution and punishment, an enabling rule of grammar or of logical calculus functions independently of norm-givers and, correlatively, the agents who follow these rules are not norm-subjects. For example, when one learns the rules of a game, say basketball, one learns certain ways of moving in the field and at the same time also the fact that certain steps are allowed while others are prohibited.

---

[9] Von Wright mentions (criminal) laws as examples of prescriptions but this does not entail that he would argue that judicial systems would be merely or primarily prescriptive.

These are neither prescriptions by some agent nor conditional norms that dictate that if one wants to play basketball then one needs to do this and that. Rather what we have are limiting conditions of the activity to be learned. Referees guard against steps, accidental or deliberate, that do not belong to the game, not against actions that are part of the general behavioral repertoire of humans but have been banned by a decision of an authority, individual or collective.

Or take another example, one in which enabling norms are lost. A demented person may lose her basic linguistic capacity and the mastery of the enabling norms of language, such as the syntactical rules that define the sentential and clausal structures of (the) language. When this happens, she may be subjected to diverse forms of therapeutic treatments, pharmaco-, occupational or other. These treatments are not sanctions or forms of punishment that would coerce the person to behave in the way willed by some authority but are treatments meant to help her regain the condition in which she may relearn to master the rules of language that allow her to form sentences, or at least some of these rules. We may of course say that the enabling norms of language or those of logical calculi "prescribe" our speaking and thinking, but this is a metaphorical parlance and depends in its sense on the primary usage of "prescribe" in which definite norm-authorities prescribe behaviors or types of behavior to particular norm-subjects.

Von Wright's distinction between enabling rules and prescriptions thus draws attention to a habit of thought that tends to dominate contemporary discussions concerning norms and normativity, in many political, pedagogical and moral contexts. We easily take all norms to be prescriptions or essentially similar to prescriptions. More particularly, we tend to assume that most norms and rules operate like traditional laws, that is, by sanctions and punishments. When no clear penalties can be identified, we easily construe quasi-sanctions in our thoughts and end up presenting, for example, various psychological and social harms as such.

4. Directives

The third category that von Wright distinguishes is the category of directives (also: technical norms). These are norms with means-ends structures, that is, norms that operate as determining means for the sake of attaining certain willed ends: "They presuppose ends of human action and necessary relationships of acts to these ends" (von Wright 1963a, p. 15). Examples of such norms include directives of use for medicines, instruments and technical devices.

The distinguishing feature of directives is that they have the logical form of the conditional. Accordingly, their standard formulation is an if-then sentence in which the antecedent specifies a willed or wanted thing and the consequent specifies what must, what has to or what ought to be done (or not) for the wanted thing to be achieved. Von Wright emphasizes this structural feature and argues that directives are logically neither prescriptive nor descriptive but exactly conditional and involve a specification of a volition in their antecedent (von Wright 1963a, p. 10; cf. 1963b, p. 160ff.). As such, these norms differ from both enabling rules and prescriptions which both are categorical in their logical form.

This must not be taken to imply that all conditionals would be directives or involve such norms. Von Wright draws attention to two other categories of conditional norms neither of which involves a volition (von Wright 1963a, pp. 10, 101). He calls "hypothetical" the norms that are formulated by if-then sentences but do not involve specification of a willed goal and a means to such a goal. Rather than concerning volitional aims, hypothetical norms concern contingent circumstances and order what should be done if a certain contingency arises. An example of such hypothetical behavioural norms would be the preventive measure "If the dog barks do not run!" or the instruction "You must tell me, if you hear any rumours about them". The third type of conditional norms are called "anankastic". These are norms that state necessary conditions between goals and actions, independently of anybody's volitions. An example of such a conditional is "If the house is to be made habitable it ought to be heated" (von Wright 1963a, p. 10; cf. 1963b, p. 158ff.).

Thus defined, directives and prescriptions depend on the human will: prescriptions issue from someone's will and directives involve voluntarily posited goals. This distinguishes both from enabling or constitutive rules that do not depend on anybody's will or volition. On the other hand, prescriptions and enabling rules are both categorical in von Wright's analysis, lacking the conditional character of directives (cf. Kelsen [1925] 1991, pp. 349–350).

Von Wright argues that his tripartite distinction between enabling norms, prescriptive norms and directives covers the main senses of normativity operative in philosophical discussion (von Wright 1963a, pp. 15–16). These three senses are primitive in that they cannot be defined by one another and by any further sense of "norm". However, von Wright identifies three additional senses of the term "norm" that combine elements of these basic ones but also add new elements. These additional types of norms are: (iv) customs (also: social habits), (v) moral principles and (vi) ideal principles. I will illuminate their specific characteristics in the following three sections.

5. Customs

Customs or social habits have certain similarities with both enabling rules and prescriptions, but they cannot be categorized as either since they also differ from both in important respects. Examples of such norms include the cultural norms of greeting, eating, dressing, marrying and burying the dead. These vary across lived space and time.

Customs are similar to prescriptions in directing or "regulating" behaviors and influencing the conducts of both individuals and groups. They can be said to be "acquired" by whole communities and "imposed" on their individual members. Thus, customs like prescriptions exert "normative pressure" on individuals.[10] Moreover, they involve various

[10] Another important similarity between customs and prescriptions is that both differ from the regularities of nature in respect to deviations. Individual members can violate the customs of the community and "break" its habits in an analogous way as they can break the law. In contrast, deterministic "laws" of physics do not allow any exceptions, and statistical "laws" of nature that do allow exceptional occurrences are not violated by exceptions.

punitive measures "whereby the community reacts to those members who do not conform to its customs" (von Wright 1963a, 9).[11]

Despite these similarities, von Wright argues, customs and prescriptions also differ in crucial respects: most importantly, customs lack authoritative source and promulgation methods and processes (von Wright 1963a, pp. 9, 25; 1963b, pp. 158–159). On this basis, von Wright argues that even if customs may be characterized as "anonymous and implicit prescriptions" on the basis of their similarities with genuine prescriptions, they should not be reified or mystified by suggesting that the historical communities or cultures that acquire them and impose them on their members would be norm-authorities:

> [T]here are important differences [...] between customs and prescriptions. Customs first of all are not given by any authority to subjects. If we can speak of an authority behind the custom at all this authority would be the community itself including both its past and its present members. Customs could aptly be characterized as anonymous norms or prescriptions. But this characterization must not encourage any mysticism about the community as a norm-giver (von Wright 1963a, pp. 17–18).

The second, and related, difference is that even though deviations from customs may be "punished" by social marginalization and exclusion (cf. Doyon & Breyer 2015, p. 1), such punishments are very different from punishments by law or prescriptions more generally, both in terms of content and in terms of execution and its conditions. In the former case, the norm-breaker is shunned, ostracized, marginalized and/or "cast out" from most or all social interaction; in the second case, she is subjected to a specific physical or economic restriction, such as fining, imprisonment, dismemberment or death (von Wright 1963a, p. 9). Again, one should not blur the difference between the two types of norms by talking

---

[11] This difference was obscured in 20th century political analyses which, inspired by Foucault's and Deleuze's analyses of power, paralleled political projects and subversive actions with mutations and transformations of vital forces (e.g. Butler 1990, p. 145; Braidotti 2002, p. 134ff.; cf. Rand 2011).

about social "punishments" as retributions or about the social outcast as an outlaw.

Customs are similar to enabling rules in determining complete ways of acting and living characteristic of communities. They are also similar in having thick histories and dynamic ways of changing and developing. These similarities should not lead us to overlook an essential difference in the manners in which these two types of norms determine their characteristic activities. This difference concerns the coverage of the norm: Whereas the enabling rules of playing, speaking or thinking define all possible "moves" of these activities, customs merely differentiate between approved and non-approved behaviors, both of which remain possible within the community. Greeting, for example, is highly customary in most communities and violations against this norm are often strongly disapproved. However, a person who never greets anybody is both thinkable and tolerable in all such communities. Eventually such a person may become an outcast, but this status does not make her an outlaw nor a mere onlooker of social exchanges. By using the metaphor of gaming, we can say that the "moves" of such a person are highly unusual and also disapproved by the other "players" of "the social game", but despite their exceptional and condemned character they are still moves of the game, moves countered by other moves and moves responded to by other players.[12]

In summary, "customs resemble [enabling] rules in that they determine or quasi-define certain patterns of conduct – and prescriptions in that they exert a 'normative pressure' on the members of a community to conform to these patterns" (von Wright 1963a, p. 16). They differ, however, crucially from enabling rules in the manner in which they determine behavior and from prescriptions in the type of pressure that they exercise.

---

[12] Émile Durkheim's discussion of rules and norms in his *The Division of Labor in Society* (*De la division du travail social*, 1893) serves as an example of overambitious theoretization which models all norm types on one specific type, in this case customs (and prescriptions) (p. 121ff.).

6. Moral principles

Moral principles, which are also called "moral rules" and "moral laws", are the second additional category of norms that von Wright distinguishes. The paradigmatic example of a moral norm is the one that obliges us to keep our promises. Other examples include the norm according to which children ought to honor their parents, the principle that the innocent should not be punished, and the golden rule according to which one should love one's neighbor as one loves oneself. The coverages of such principles vary greatly. Some are highly context-dependent while others are absolute. For example, the moral rules of sexual ethics seem to depend greatly on contextual factors, whereas the commands "Do not kill!" and "Respect the elderly!" have an absolute character.

Von Wright draws attention to the fact that moral principles are traditionally and still often taken to be either a subspecies of prescriptions or else a subspecies of directives. The third alternative, common in philosophical literature, is to argue that these norms are *sui generis*. Von Wright rejects all three analyses and argues that in truth moral principles do not belong to any logically distinct category of norms nor do they form a category of their own. Instead, they combine features of several other types of norms, most importantly features of enabling norms, prescriptions and customs. And not only this, but more: different cases of moral principles combine logically different normative features and do this in several different ways. Thus, the category of moral principles is logically heterogenous in von Wright's analysis:

> The peculiarity of moral norms as I see them is not that they form an autonomous group of their own; it is rather that they have complicated logical affinities to the other main types of norms and to the value-notion of good and evil. To understand the nature of moral norms is therefore not to discover some unique feature in them [structural or other]; it is to survey their complex relationships to a number of other things (von Wright 1963a, p. 13).[13]

[13] On this basis, von Wright attacks both theistic, utilitarian and eudaemonistic ethics by arguing that moral norms are neither prescriptions by authorities nor conditional directives defined by willed goals. Moreover,

Some moral principles are, according to von Wright, similar to the enabling rules of language and logical calculi in constituting complete forms of action while others are more like customs in being contextual or situational. For example, the obligation to keep one's promise is similar to an enabling rule of grammar in defining the institution of giving and taking promises. In contrast, sexual ethics includes custom-like moral principles, such as the norms against inbreeding and the norms against zoophilia and bestiality (von Wright 1963a, 12).[14]

Neither are all moral principles prescriptions, heteronomous or autonomous. Prescriptions have an important role to play in moral education and in the expression of morality, but this should not be taken to imply that all moral principles would have the logical character of prescriptions (Wright 1963a, 12–13). In this context, von Wright also argues that self-regulating "commands" must be kept separate from authority-based prescriptions: "Such 'autonomous' prescriptions given by man to himself are [...] very unlike the 'heteronomous' prescriptions, categorical or hypothetical, given by a norm-authority to some norm-subjects. It is doubtful whether one should call the former 'prescriptions' at all" (von Wright 1963a, p. 11). This means that philosophical discussions on self-regulation must not be modelled on authority-based regulations by others.

On the basis of this analysis, von Wright attacks similarly theistic, utilitarian and eudaemonistic ethics, arguing that moral norms are neither prescriptions by authorities nor conditional technical norms defined by willed goals. Moreover,

---

he also rejects deontic theories by arguing that moral norms are not *sui generis*. Ultimately, von Wright contends that the logical heterogeneity of moral norms cannot be philosophically elucidated without a profound analysis of the various senses of goodness and the good, that is, the sense of value. What turns out to be logically foundational is the concept of the goodness of human being. Thus, as pointed out above, von Wright's analyses in *Norm and Action* point to *The Varieties of Goodness* (1963).

[14] Von Wright seems to be informed here by the work of another Finnish philosopher Edward Westermarck who studied the cultural and historical nature of various moral norms in his widely influential anthropological treatises, most importantly *The History of Human Marriage* (1890) and *The Origin and Development of Moral Ideas* (1906).

he also rejects deontic theories on the basis that moral norms are not *sui generis* either. Ultimately, he then contends that the logical heterogeneity of moral norms cannot be philosophically elucidated without a profound analysis of the various senses of goodness and the good, that is, the sense of value. What turns out to be logically foundational is the concept of the goodness of human being. Thus, von Wright's analyses in *Norm and Action* point to his parallel work on *The Varieties of Goodness* (1963).

7. Ideal principles

The final sense in which we speak about norms, distinguished by von Wright, is the sense of ideal principles. These are not norms of doing, acting or behaving but are norms of *being*. Here von Wright draws directly from the phenomenologist Max Scheler's distinctions between "*Tunsollen*" and "*Seinsollen*" and between "normative ought" (*normatives Sollen*) and "ideal ought" (*ideales Sollen*) and argues that the normativity of doing and that of being must be kept distinct, since the former implies the concept of *rule-following* while the latter suggests that adherence to norms is like *seeking something* or constantly *striving for something* (von Wright 1963a, p. 15).[15]

Examples of ideal principles include norms that articulate ethical virtues and norms that govern professional excellences. When we state, for instance, that a human person ought to be generous, truthful or just, we express, in von Wright's analysis, a moral norm that has the logical character of an ideal principle; and when we strive for these perfections, then we act according to an ideal principle which is part of morality. Analogously, when one points out that a university professor should be patient with her students but at the same time also firm, then one states an ideal principle of teaching. And if one then acts with firmness and patience in the class room, despite possible complaints, then one is striving for

[15] When making the distinction between norms as rules and prescriptions and norms as ideal principles, von Wright refers explicitly to Scheler's *Der Formalismus in der Ethik und die materiale Wertethik* (1913–1916) but also to Nicolai Hartmann's *Ethik* (1926) and G.E. Moore's "The nature of moral philosophy" (1922) (von Wright 1963a, pp. 14–22).

professional excellence, not accommodating oneself to traditional academic conventions or following the latest pedagogical directives issued from the administration.

Such striving or seeking to be good at something (teaching or leading a research community) may occasionally demand that one questions or abandons the traditional norms and standards that regulate the practicing of the profession in one's own community. More radically, sometimes one may also need to act against the general conventions that regulate the profession globally across communities and even whole cultures, contemporary and past. In this way, one may become a social outcast but that does not imply that one would have lost one's task and obligation as a professional. Instead, one may act as a reformer or revolutionary. Ideal principles thus differ from norms understood as customs or social habits: in the case of such norms, one's responsibility is primarily for the ideal, not to the fellow practitioners of the profession (cf. Frankfurt 2004; Heinämaa 2014; Melle 2007; 2002).

In having this goal-oriented character, ideal principles may seem similar to directives. Von Wright argues, however, that we must not make the logical mistake of confusing our striving for professional, epistemic or moral ideals with the processes in which we follow directives and try to achieve goals. This is because ideal principles resemble enabling or constitutive norms: they are not motivational causes for our actions but are conditions that define ways of being.

> There is a certain similarity between ideal rules and [directives]. Striving for the ideal resembles the pursuit of an end. It would however be a mistake to think of the ideal rules as norms concerning means to ends. In order to be a good teacher a man ought to have such and such qualities. [...] But those qualities of a man which determine his goodness as a teacher are not causally related to the ideal – as the use of ladder may be a causal prerequisite of fetching a book from a shelf. The former relation is conceptual (logical). The ideal rules determine a concept, e.g. the

> concept of a (good) teacher or soldier. In this they are similar to rules of a game (von Wright 1963a, p. 15).[16]

For example, let us assume that the ideal principle of being a teacher includes the task (regulative idea) of being both firm and kind. If this holds, then a teacher must address and treat her students firmly and kindly. This must not be construed as a conditional that dictates that if someone settles to be a teacher, then she must act firmly and kindly toward students. Rather than specifying causally, functionally or motivationally what the person needs to do in order to figure or operate as a teacher, the ideal principle defines what it entails to be a teacher.

By defining ways of being, ideal principles govern actions and types of action but also *modes* of acting (cf. Audi 2016). Whatever the soldier does, he is obliged to act bravely and in a disciplined manner; analogously, firmness and kindness should characterize the attitude of a teacher in her various activities; and finally, the virtuous person with moral integrity is expected to act justly and honestly in all her dealings (cf. von Wright 1963b, p. 139). An academic who addresses his colleagues respectfully but criticizes his students dismissingly and scornfully is not a good teacher (or scholar) and, depending on the gravity of her manners, may not be a teacher (or scholar) at all.

8. Epilogue: phenomenological perspective

As pointed out above, in his discussion of ideal principles, von Wright refers to the analyses of two early phenomenologists, Max Scheler and Nicolai Hartmann. Both contribute to the critical discussion of normativity by offering axiological analysis of the relations between rules of action, on the one hand, and values, on the other hand. Both argue that values as such do not regulate actions. Values are ideal principles of being and as such they merely determine what ought to *be* without commanding or dictating what must be *done*. This "normative inertness" of values in respect to action is due to

---

[16] Von Wright also uses the term "ideal rules" for ideal principles. He decides this on the basis that ideal principles are similar to enabling or constitutive rules in delineating and defining complete forms of activities.

their structure: they do not contain in themselves any reference to volitions or acts of willing, integral to action.[17] The value of beauty, for example, "demands" that beauty ought to be, but it does not thereby command any beautiful actions or actions approximating or pursuing beauty. Only when considered by a willing and reflective subject, Scheler and Hartmann contend, can an ideal value mobilize and direct action (Scheler 1913–1916, pp. 187–188, 214; Hartmann 1926, pp. 154–159, 171–172; cf. Hessen 1958, pp. 83–84; Kelly 2011, pp. 110–112).

Von Wright agrees on this basic point. In the chapter titled "'Good' and 'Must'" in *The Varieties of Goodness*, he attacks the common notion according to which the concepts of value and goodness are intrinsically normative (von Wright 1963b, p. 155ff.). His account is similar to Scheler's and Hartmann's in suggesting that action can become value-directed but only on the basis of volition and its imperatives: "I tend to think that it is only the aspect of norms as practical necessities […)] which bears an intrinsic relationship to ideas of the good. Other aspects of the normative may become value-oriented only through the intermediary of the [practically necessitating] aspect" (von Wright 1963b, p. 177).

The main divergence between von Wright and his phenomenological predecessors is in questions concerning the dependency relations between norms and values. Both Scheler and Harman explicitly contend that the norms of moral action (obligations) are founded on moral values (e.g. Scheler 1913–1916, p. 187; Hartman 1926, p. 159). Von Wright, by contrast, maintains that the question concerning the foundational relations between norms and values should not be decided on the basis of analyses focused on moral norms and values exclusively. Rather, the task is to widen the perspective and also

[17] In Hartmann's analysis, a value essentially involves an obligation of being (*Seinsollen*) (Hartmann 1926, pp. 154–156). Scheler, in contrast, contends that as ideal objectivities, values do not involve any obliging moments whatsoever, neither obligations of being nor obligations of doing. They become obliging, however, when they are considered in relation to a possible reality. This as such does not, in Scheler's analysis, transform them to norms of doing or acting. What is needed for such a modification, he agrees with Hartman, is a reference to striving or willing (Scheler 1913–1916, p. 187).

inquire into the relations between diverse extra-moral norms and values:

> It seems to me that the discussion of the relations between norms and values even in recent times has suffered from the narrowing and obscuring implications of the term 'moral'. If we want to get to know what values as such have to do with norms as such or to know the general nature of the connection, if there is one, between norms and values, we must disentangle the two from their associations with morality and study them in the widest possible generality (von Wright 1963b, p. 156).

The phenomenological tradition offers original insights that promote this course of investigation. It entails an argument about epistemic norms that parallels Scheler's and Hartman's arguments about moral norms. We find this argument in Edmund Husserl's *Formal and Transcendental Logic* (1927). In this work, Husserl contends that sciences in general are regulated not by mere norms of reasoning but more fundamentally by ideal principles of being. The logical laws that guide the scientists do not operate by ruling over their activities of inferring, proving or arguing but by conditioning the ideal structures and formations (*Gebilde*) that they aim at, that is, evident and true concepts, judgments, theories and, ultimately, sciences themselves as a complicated system of such formations (Husserl [1927] 1974, pp. 228–230/258–260; cf. Hartimo 2018; 2019; 2020).

University of Jyväskylä

References

Audi, R. (2016), *Means, Ends and Persons: The Meaning and Psychological Dimensions of Kant's Humanity Formula*, Oxford University Press, Oxford.

Braidotti, R. (2002), *Metamorphoses: Toward a Materialist Theory of Becoming*, Polity Press, Cambridge UK, Malden USA.

Butler, J. (1990), *Gender Trouble: Feminism and the Subversion of Identity*, Routledge, New York.

Doyon, M. & T. Breyer (2015), "Introduction", in *Normativity in Perception*, eds. M. Doyon & T. Breyer, Palgrave Macmillan, New York, pp. 1–13.

Durkheim, É. [1893] 2013: *The Division of Labor in Society*, ed. S. Lukes, Second edition, trans. W. D. Halls, Palgrave, Mcmillan, London.

Finlay, S. (2019), "Defining normativity", in *Dimensions of Normativity: New Essays on Metaethics and Jurisprudence*, eds. D. Plunkett, S. Saphiro, and K. Toh, Oxford University Press, Oxford.

Hartimo, M. (2018), "Radical *Besinnung* in *Formale und transzendentale Logik* (1929)", *Husserl Studies*, 34, pp. 247–266.

Hartimo, M. (2019), "Husserl on Kant and the critical view of logic," *Inquiry: An International Journal of Philosophy*, DOI: 10.1080/0020174X.2019.1651089

Hartimo, M. (2020), *Husserl and Mathematics*, Cambridge University Press, Cambridge, forthcoming.

Hartmann, N. (1926), *Ethik*, Walter de Gruyter & Co, Berlin.

Heinämaa, S. (2014), "Husserl's ethics of renewal: A personalistic approach", in *New Perspectives to Aristotelianism and Its Critics*, eds. M. Tuominen, S. Heinämaa, and V. Mäkinen, Brill, Leiden and Boston, pp. 196–212.

Heinämaa, S. (2019), "Constitutive, prescriptive, technical or ideal? On the ambiguity of the term 'norm'", in *Normativity, Meaning, the Promise of Phenomenology*, eds. M. Burch, J. E. Marsh and I. McMullin, Routledge, London, 2019, pp. 9–28.

Hessen, J. (1958), *Ethik: Grundzüge einer personalistischen Wertethik*, Brill, Leiden.

Husserl, E. [1927] 1974, *Formale und transzendentale Logik: Versuch einer Kritik der logischen Vernunft*, *Husserliana Band XVII*, ed. P. Janssen, Springer, Dordrecht. In English *Formal and Transcendental Logic*, trans. D. Cairns, Martinus Nijhoff, The Hague, 1969.

Jakola, L. (2020), "Wittgenstein and G. H. von Wright's path to *The Varieties of Goodness* (1963)", *Nordic Wittgenstein Review* 9, pp. 1–14.

Kelly, E. (2011), *Material Ethics of Value: Max Scheler and Nicolai Hartmann*, Springer, Dordrecht.

Kelsen, H. ([1925] 1991). *General Theory of Norms*, trans. M. Harney, Clarendon Press, Oxford.

Lorini G. & W. Zelaniec (2018), "The background of constitutive norms: Introduction", *Argumenta* 4, pp. 9–19.

Melle, U. (2002), "Edmund Husserl: From reason to love," in *The Phenomenological Approaches to Moral Philosophy*, eds. J. Drummond and L. Embree, Dordrecht, Boston, London, Kluwer Academic Press.

Melle, U. (2007), "Husserl's personalistic ethics", *Husserl Studies* 23, pp. 1–15.

Mohanty, J. N. (1997), *Phenomenology: Between Essentialism and Transcendental Philosophy*, Northwestern University Press, Evanston, Illinois.

Parfit, D. (2011), *On What Matters, Volume 2,* Oxford University Press, Oxford.

Passerini, L. & P. Di Lucia (eds.) (2017), *Norm: What Is It? Ontological and Pragmatical Perspectives*, Special Issue of *Phenomenology and Mind*.

Rand, S. (2011), "Organism, normativity, plasticity: Canguilhem, Kant, Malabou", *Continental Philosophy Review* 44, pp. 341–357.

Scheler, M. (1913–1916), *Der Formalismus in der Ethik und die materiale Wertethik*, Verlag von Max Niemeyer, Halle.

Venturinha, N. (2020), "Wittgenstein and the method of von Wright's *Varieties of Goodness*", *Philosophical Investigations*, https://doi.org/10.1111/phin.12281

Wittgenstein, L. ([1953] 1997), *Philosophische Untersuchungen/Philosophical Investigations, Second Edition*, trans. G. E. M. Anscombe, Blackwell Publishers, Malden and Oxford.

Wittgenstein, L. ([1967] 1970), *Zettel*, eds. G. E. M. Anscombe and G. H. von Wright, trans. G. E. M. Anscombe, California, University of California Press, Berkeley and Los Angeles.

Wittgenstein, L. ([1977] 1980), *Culture and Value*, University of Chicago Press, Chicago.

von Wright, G. H. (1951), "Deontic logic", *Mind* LX, pp. 1–15.

von Wright, G. H. (1963a), *Norm and Action*, Routledge & Kegan Paul, London.

von Wright, G. H. (1963b), *The Varieties of Goodness*, Routledge & Kegan Paul, London.

How Is Normativity Possible? A Holistic-Pragmatist Perspective

SAMI PIHLSTRÖM

Introduction

The philosophical question concerning the very possibility of normativity captures an essential puzzlement about what it is or means to be human – inherited into our contemporary discussions all the way from Plato's or Socrates's criticism of Thrasymachus in Book I of *The Republic* – and in my view it does so more deeply than, for example, the question concerning the nature of consciousness. The problem is as old as the Greeks' distinction between *fysis* and *nomos*; however, it is not solved by claiming that normativity is based on, or arises from, human convention, social negotiation, or something similar. That is merely to restate the issue, because such conventional practices already presuppose a normative context. We are thus dealing with a problem of infinite regress or circularity, while on the other hand we may also ask whether the relevant kind of circularity is vicious or perhaps rather beneficial. In fact, recognizing its *inevitability* is part of my pragmatist reaction to our issue. In a sense, the question about the possibility of normativity may have no "solution" at all; what needs to be done is learning to live with it.

Focusing on the idea that a normative context seems to be already presupposed in attempts to account for its possibility, this paper starts from the conviction that we need to develop a *transcendental* inquiry into normativity. Such inquiries include, in my somewhat relaxed sense, not only Immanuel Kant's theory of the categories of the understanding as well

as the moral law,[1] but also classical pragmatists' like Charles S. Peirce's and William James's views on habits of action, human practices, and constructive purposive activities, as well as the later Wittgenstein's philosophy of language-games, forms of life, and rule-following.[2] These transcendental conceptions of normativity are to be distinguished from metaphysical theories of the grounds of normativity based on, for example, emergence as well as naturalizing reductions, which seem to try to account for normativity from an "external" rather than reflexively "internal" perspective (but cf. Pihlström 2010). In brief, a transcendental philosophy of normativity seeks to understand and further articulate our commitment to normativity *from within* a framework (practice, form of life) already defined by such a commitment. It may thus seek to offer a transcendental "deduction" in the Kantian sense of rendering our commitment to norms *legitimate*[3] – rather than a metaphysical, scientific, or empirical explanation of how or why the norms we do commit to have arisen.

For the same reason, transcendental investigations of normativity also need to be distinguished from the mainstream approaches of social ontology. A key concept often employed in social ontology that we will, however, examine in some more detail in this context is *recognition*. I think of contemporary recognition theory as lying somewhere between transcendental and non-transcendental approaches to

---

[1] Indeed, normativity is at the core of Kant's projects both in theoretical and in practical philosophy.

[2] I am not including Hegel in this camp, as I believe the transcendental philosopher of normativity should follow Kant rather than Hegel, but I will comment on the concept of recognition, a concept with a strongly Hegelian history, in what follows. While I won't be able to discuss their work in this paper, my approach comes closer to the broadly Kantian analysis of the "sources of normativity" by Korsgaard (1996) than, say, Taylor's (1992) more Hegelian position. (For my brief reading of Korsgaard's project and her notion of practical identity, see Pihlström 2005, chapter 3.) On the other hand, Taylor's (1989) notion of "strong evaluation" is readily comparable to the idea of the transcendental constitution of irreducible normativity.

[3] The analogy, of course, is to the *de jure* question Kant poses in his transcendental deduction of the categories of the understanding.

normativity. While recognition is still *contingent* in a way a fully transcendental ground of normativity cannot be (or so I will argue), it can be claimed to be *constitutive* of social facts and institutions, or even human personhood. From the perspective of the present inquiry, an essential question is whether the relevant kind of constitutivity is metaphysical in a non-transcendental sense or transcendental in a (quasi-) Kantian sense. I am not going to examine Kant's own views here, but we should recall the idea, strongly albeit somewhat implicitly present in the First Critique, that the categories of the understanding are constitutive of all humanly possible experience and its objects by providing *normative requirements for what it is to be an object for us*. I will try to explain why I am not convinced that the kind of normativity that recognition brings to our social world operates at the same transcendentally constitutive level.[4]

I have elsewhere recommended a *negative method* for various philosophical purposes of pragmatically elucidating what certain concepts mean for us in our lives (cf., e.g., Pihlström 2014); in the present case, such a method would urge us to take a serious look at various (actual and possible) violations, eliminations, or reductions of normativity in order to understand and appreciate what normativity (positively) is. These may include, for instance, reductionisms of various stripes (e.g., attempts to reduce humanly distinctive normativity to, say, brain activity or evolutionary processes, or both)[5] or

[4] As a first approximation of our main issues, it might be noted that the United Nations Declaration of Human Rights mentions, in §6, the right to recognition before the law – something comparable to what Hannah Arendt called "the right to have rights" (cf. Bernstein 2018). Such rights, I will argue, cannot be the *source* of normativity, as they are only possible in a context that is already normatively structured. These fundamental philosophical issues become strikingly practical as soon as one notes that, for example, the situation of former Isis women and children at refugee camps or of the Guantanamo prisoners can be seen as a state of "rightlessness": these unfortunate people, for various reasons (for which they are not innocent), have ended up in circumstances in which they do not seem to be recognized by any normative system, or to have even the right to have (e.g., legal) rights.

[5] Analogous criticism should be directed at attempts to reduce political normativity to something non-political; the autonomy of the political –

populist politicians' (e.g., Donald Trump's) tendency to step out of normative contracts in international relations, preferring something like a Thrasymachian politics of force. Such tendencies do not acknowledge what is distinctive in the human form of life, because to be human is to live in a normative "space of reasons" that cannot be simply replaced by non-normative structures. For the same reason, *human dignity* – a notion I will return to in due course – cannot be grounded in recognition, or any other contingent attitudes, just as morality cannot be reduced to brain activity, for instance.

A philosophical analysis on normativity is more widely relevant than it might initially appear. Its significance ranges from daily phenomena of interpersonal encounters to extremely complex political processes, and beyond. A "negative" investigation of normativity involves, moreover, a *self-criticism* of our normative form of life parallel to Kant's analysis of the illegitimate transgressions of human reason manifested in the transcendental illusions he analyzes in the Transcendental Dialectic of the First Critique. For example, what I am calling "naturalizing reductions" of normativity can be seen as analogous to such illusions. There may even seem to be a kind of unavoidability inherent in them comparable to the unavoidability of transcendental illusion: it might seem that norms just *have to* be grounded in natural facts, even though in a sense they *cannot* be, more or less like it might in the Kantian context seem to our reason both that, for instance, the world as a totality must have spatio-temporal limits and that the world cannot have such limits.[6] Naturalizing reductions tend to replace the philosophical (transcendental) question about the very possibility of normativity by an empirical and/or causal explanatory question about the emergence and development of normativity, and while there is of course nothing wrong with the latter kind of question as

---

and more generally normative – sphere is to be acknowledged. For example, when radically right-wing "ethno-nationalist" populists emphasize the genetic similarity among those belonging to an *ethnos* (e.g., the Finns), they are reducing the normative (nationality) to the natural (genes).

[6] No reading of Kant's account of the antinomies, or other transcendental illusions, is of course attempted here.

such, this replacement tendency leads us seriously astray – or so I will argue.

"Human nature": normativity as a philosophical-anthropological and cultural issue

The issue we are exploring goes back to the problem of "human nature", or *philosophical anthropology*.[7] In terms of Heikki Kannisto's (1984) useful fourfold classification of the "ideal types" of philosophical conceptions of humanity, we may (in a simplified way) pose our basic question in this form: do we as human beings belong to an objective, independent, cosmic normative order (*essentialism*), are we without any such order (*naturalism*, *existentialism*), or are we creators of our own cultural normative order (*culturalism*)? How, moreover, can we decide between these alternatives?

Following Kannisto's terminology, we may say that reductive naturalism "factualizes" any normative order we might take ourselves to be inhabiting by reducing the classical essentialists' (e.g., Aristotle's) postulation of cosmic teleological normativity into mere nature, i.e., contingent and fully natural matters of fact, and thereby moves human beings out of any distinctive normative space of reasons to the realm of natural law. Such naturalism may be argued to be problematic precisely because of its inability to account for genuine normativity, but on the other hand it has at least since the Enlightenment plausibly questioned the classical essentialist postulation of Platonic or Aristotelian cosmic normativity beyond our concrete and contingent human activities (as well as Christian or other theological variants thereof).[8] In contrast to both naturalism and essentialism, culturalism may be ar-

---

[7] Cf. Kannisto 1984; Pihlström 2003, 2016; and especially, for an indication of the current recovery of philosophical anthropology, particularly in relation to debates over naturalism, Honenberger 2016.

[8] Another line of argument critical of classical essentialism is of course existentialism, according to which human beings have no ahistorical metaphysical essence but individually create their own lives and normative principles in the contingent (absurd) situations they happen to find themselves in. Due to its radical individualism, existentialism might also lead to a fragmentation of normativity, though for reasons different from naturalism.

gued to be a plausible way of accounting for normativity: our normative sphere is humanly constructed; it is, for us, fully real without emanating from any Platonic or other transcendent sources beyond our human forms of life.[9]

However, there is a problem analogous to naturalism within culturalism itself, because, ironically, *cultural relativism*, an arguably natural articulation or development of culturalism, may be just another way of "refactualizing" the normative order into mere contingent matters of fact.[10] While culturalism emphasizes that human beings live in a normatively structured human world that is largely of their own making, rather than being placed within a pre-established teleology and cosmic normativity in a classical (e.g., Aristotelian) sense, this idea rapidly collapses into relativism as soon as we admit that any such structuring of normative frameworks takes place within specific and spatio-temporally localized historical cultural spheres.[11] The challenge for culturalist philosophical anthropology is to maintain as much irreducible normativity as possible without postulating any Platonic or Aristotelian essentialist normativity that cannot be grounded in natural processes. This is, I suppose, the traditional issue of nature vs. culture all over again, with broadly culturalist approaches ranging from Kant's distinction between theoretical and practical reason (with their specific normative tasks) to John Dewey's (1929) analysis of "experience and nature" and Wittgenstein's (1953) claim to study "the natural history of human forms of life".

It is, in my view, exactly this challenge that *pragmatism* – speaking of pragmatism generally as a philosophical orientation from Peirce and James to Rorty and Putnam, and beyond

---

[9] This can be regarded as, essentially, a Kantian-cum-Wittgensteinian framework for philosophical anthropology, with the world-constituting activity of the Kantian transcendental subject reconceptualized as a Wittgensteinian normatively structured form of life.

[10] My worries here concern only the most radical forms of relativism. There are certainly responsible moderate forms around.

[11] Hilary Putnam (1983) also aptly suggests that cultural relativism ought to be understood as a species of (reductive) naturalism. Both are, in short, haunted by the loss of normativity.

– has seriously aspired to meet.[12] I have always found pragmatism one of the most promising philosophical approaches in this discussion, as it takes seriously both non-reductive naturalism and irreducible cultural normativity.[13] Instead of pragmatism, I will here shortly turn to the concept of recognition, however (postponing a brief discussion of pragmatism to a later section). Could recognition theory, we may now ask, also be employed to make sense of the emergence of the normative order as such? Or does it already presuppose a normative order? Is there a "first" recognition act upon which the normativity of our social world could in principle be based? These are among the questions that need to be addressed by anyone taking seriously the task of bridging the gap between naturalism and culturalism, and it might be tempting to think that recognition theory could resolve this issue.

Without claiming that contemporary recognition theorists are actually in the business of doing so, it might be tempting to see recognition as a way of accounting for the possibility of normativity. Thus, normativity would be grounded in acts of recognition. (In a sense this goes back to Hegel's dialectic of the master and the slave.) But if this is the case, is there an "original" – and hence natural – situation with no recognition acts in place yet? There would, ideally, have to be, if recognition were to offer a ground for normativity in the sense of turning initially natural facts into normative statuses. If so, then *how* does, or how did, normativity emerge from such a purely factual situation? From recognitions of normative statuses perhaps? But then how do we know (or how did the "first recognizers" know) to whom, or to what, such recogni-

---

[12] I try to offer a pragmatist yet transcendental philosophical anthropology (with special emphasis on the problem of death and mortality) in Pihlström 2016.

[13] More generally, the three *critical* (transcendental) philosophies of normativity that I am trying to understand and develop further in my own work (e.g., Pihlström 2016) are Kantianism, pragmatism, and Wittgensteinianism. Obviously they cannot be explored here at any length, but this essay as a whole is crucially informed by a pragmatist approach to normativity (as will become more explicit in due course). For a more comprehensive argument for a pragmatist articulation of transcendental philosophy (which I recognize as controversial), see, e.g., Pihlström 2003.

tion acts should, or even could, be directed? Let me re-emphasize that I find these much more important – and more human – questions than the allegedly deep question of how, say, consciousness emerged, or emerges, from non-conscious matter. The questions concerning recognition and normativity are presumably also less prone to lead to postulations of mysterious qualia or other strange non-natural entities that may not seem to fit into the scientific worldview.[14] The Kantian issue of legitimacy, or entitlement, cannot be settled by focusing on any quasi-scientific factual question.

Let me illustrate our problem with reference to the very distinctive horrors brought into our social and cultural world by Nazism. What is relevant here is, arguably, the *Nazis' destruction of (almost) all normative* (ethical, political, legal, etc.) *statuses of the victims*, or most of them at least (cf. Snyder 2010, 2015). Ironically, the Nazis did have their own "laws" and a "*Volksgericht*" delivering "legal" judgments within their bizarre society, but these perversions of normativity were ultimately based on a thoroughly biologistic doctrine of *Lebensraum* and racism; accordingly, the normativity at work in the Nazi system was, arguably, almost entirely reducible, and indeed rather literally reduced, to both the victims' and the perpetrators' racial and biological contingencies, such as the Germans' allegedly natural need for *Lebensraum* and the fact that the non-Aryan "lower races" of the East were on the way. As Holocaust writers like Primo Levi forcefully testify, the Nazis largely succeeded in reducing their victims into mere beasts, not merely by what they concretely did to them but also by using the kind of non-humanizing language they used (Levi 1988), while in a sense remaining human themselves, because remaining guilty and responsible for what they did. This reduction of human beings to mere animals is carefully analyzed in Holocaust literature, including Levi's compelling work.[15] But it required a

[14] Putnam (1999) persuasively argues that there is something seriously wrong in the temptation to think of the mind in terms of a (quasi-) scientific mystery in the first place. I agree, though I am not investigating the mind here, except in the extremely broad sense that normativity (naturally) requires mental or psychological creatures.

[15] On Levi's importance in the acknowledgment of the meaninglessness of suffering, see Pihlström 2020, chapter 6. In a Wittgensteinian analysis of

philosophical-political analysis of the magnitude of Hannah Arendt's to show what novel kind of crime the Nazi crime was. In Arendtian terms, the elimination of human *spontaneity* in totalitarianism (Arendt 1958) can be seen as a version of the reduction of normativity into mere natural factuality, or even non-human bestiality.

Our problem is that normativity is *irreducible but not non- or supernatural*. It is crucial to avoid both "bald naturalism" and "rampant Platonism", as John McDowell (1996) aptly calls them. This needs to be done across the board from logic and epistemology to ethics and political philosophy.[16]

Recognition

Let us move on to a slightly more detailed discussion of recognition as a ground of normativity. I am not seeking to offer any comprehensive account of contemporary recognition theory; my remarks may be understood as critical suggestions that would, I think, have to be addressed by anyone

---

Levi, Sparti (2005) argues that acknowledging others as humans, or the lack thereof, needs a form of life as its context – and my argument in this paper will come close to this line of thought. However, Sparti speaks about our *responsibility* of acknowledging others; again, the question is how (and when) such a normative responsibility arises. Doesn't it already need a normative context to be so much as possible? For a compelling analysis of the way in which the horror of the Holocaust moves us beyond language and the normative, see Cavarero 2018.

[16] McDowell's (1996) notion of "second nature" might also be helpful here (cf. Pihlström 2003, 2005), but if so, it also needs, for our purposes, to be *transcendentally* (as well as pragmatically) articulated, with normativity naturally based on (but not reduced to) our on-going critical self-reflection, focusing on our constant failure to follow the norms and rules that govern our lives. (A "via negativa" method is at work here, again.) This approach might come close to Korsgaard's (1996) Kantian account of procedural normativity. Having dealt with McDowell's (and, more briefly, Korsgaard's) views on earlier occasions (Pihlström 2005, chapters 2-3), I won't dwell on this issue here, while I warmly agree with his understanding of the ethical as "a domain of rational requirements" to which we are "alerted" by acquiring appropriate conceptual capacities through enculturating upbringing (see McDowell 1996, p. 82), and with his antireductionist view that nothing non-normative can ground or justify the normative.

who proposes recognition as a "natural" (socio-)psychological ground of normativity, but this is compatible with acknowledging that contemporary recognition theorists themselves would only rarely do so.

My worry with the notion of recognition in this context, as already hinted at above, is that it may be too psychologizing and, hence, also naturalistically "factualizing" a concept to be able to account for the possibility of normativity in a sufficiently deep transcendental sense.[17] In its own way, recognition theory may seem to reduce normative structures to our *acts* of recognition, that is, something that we as contingent psychological and social individuals "naturally" do (or fail to do). As a further approximation of our problem, consider now this question: could there be a *duty* to recognize (say, someone as something) if one just doesn't "feel" the compelling demand coming from the other's point of view, such as their request for recognition, as *already* binding in any sense? The vocabulary of duties, it seems to me, would come too early here. The mere availability of such a question shows that recognition cannot be the ultimate ground of moral duty, or any duty. Or consider, again, this: if there was a *first* act of recognition, was it an idealization like the Hobbesian sovereign arising from a state of nature, or Rawlsian justice emerging from an original position behind the veil of ignorance?

---

[17] There can be no brief answer to the question (raised by one of the anonymous reviewers of this paper) why I am including pragmatism in my "relaxed" articulation of transcendental philosophy while excluding recognition theory, which might seem to have a much more intimate connection with German idealism and hence with the transcendental tradition than pragmatism does. The only obvious reason is that I see recognition theory as, *qua* Hegelian, giving up at least one basic idea of transcendental inquiry, i.e., transcendental idealism, while I see pragmatism as a Kantian approach precisely in its attempts to rearticulate transcendental idealism in a "naturalized" and historicized shape. (The same goes for the later Wittgenstein as a transcendental thinker; cf. Pihlström 2003.) Moreover, my criticism of recognition theory is restricted to the understanding of recognition as basically psychological and socio-psychological action; insofar as this perhaps overly psychologistic characterization of recognition theory is inaccurate, I am pleased to welcome recognition theory as a contribution to a (quasi-)transcendental analysis of the possibility of normativity, too.

Such an idealized postulation would in my view put the cart before the horse precisely because recognition is too contingent to account for normativity at a transcendental level, or for the grounding of the normative order in our natural psycho-social characteristics and (merely factually conceived) human nature. (The same holds, *mutatis mutandis*, for any other psychological or natural, generally non-normative, attempts to account for the grounding of normativity.)

A practice of recognition is, indeed, a *practice*. We need a well functioning set of already normatively structured and established human practices in order for there to be acts of recognition at all. Systematic recognition theory (cf., e.g., Koskinen 2017, 2019) analyzes those practices and the concepts they invoke in great philosophical detail and with admirable sophistication, but as far as I can see it cannot *ground* the normative order as such (nor is it, I suppose, necessarily taken to, though, unless recognition is proposed as *the* fundamental concept in social ontology).[18] Yet somehow norms undeniably *do* arise out of our natural ways of doing things. I would be inclined to analyze this phenomenon in terms of our "naturally" occupying or engaging in always already (for us) *irreducibly normative forms of life* (Wittgenstein), or *practices* (pragmatism). But the question remains: *how* do these forms of life or practices get their distinctive normativity? From recognition acts perhaps – but by whom, and based on what?

In a sense, recognition shares the problem of naturalism and cultural relativism: the worry is that it ultimately amounts to a "refactualization" of the normative order. It functions very well as "social glue" and is arguably *ontologically* constitutive of the social world as we know it, but it cannot *transcendentally* function as the necessary condition for the possibility of normativity. This is because there must already be a rich context of normative statuses at work in order for any act of recognition (i.e., recognizing, or failing to recognize, such statuses) to be so much as possible. This can be

[18] Again, let me emphasize that my criticism is not primarily directed at contemporary recognition theory – which might indeed have received sufficiently transcendental elaborations by its practitioners – but at a *temptation* to employ this theory in an attempt to ground normativity in contingent acts of recognition.

explicated by means of transcendental argumentation. In order for us to be able to recognize, or fail to recognize, anything whatsoever, in any sense stronger than a mere natural reaction (in principle available to "mere animals"), as having a normative status of any kind, we must already live in a normative order, a space of reasons. This is comparable to the way in which we, according to Kant, need a system of categories already in place for us to be able to have cognitive experience of any object or event – rather than a mere Humean "rhapsody" of sense impressions. Recognition can no more *ground* the possibility of normativity than the Kantian categories (as normative requirements for objecthood) can be grounded by or derived from (Humean) experience, or Wittgensteinian rules of using language within a language-game from mere marks and noises. Any theory finding recognition foundational for morality and normativity is therefore (in an extremely broad sense) "Humean" rather than Kantian.[19]

Let us elaborate on the problems and prospects of the notion of recognition in this context by taking a slightly more detailed look at a recent investigation of the topic developing and applying the original insights of Axel Honneth (2005) and other pioneers of the theory. In their introduction to the valuable new volume, *Recognition and Religion* (2019),[20] Maijastina Kahlos, Heikki J. Koskinen, and Ritva Palmén emphasize the relevance of recognition theory to the issue of normativity by reminding us that in contemporary recognition theory, recognition in its most relevant sense means that "to recognize someone is to grant another human being a positive normative status based on her personhood" (1). "On the most general level", therefore, a recognition act "means taking and treating the other *as a person*" (ibid., p. 1). When this is specified, "particular aspects of personhood" are brought into the picture, and then we can, following Honneth's seminal theory, distinguish between respect, esteem,

[19] In addition to being Hegelian, of course. There is a sense in which my discussion here parallels Kantian criticisms of Humean accounts of ethics based on sympathy. Cf. again Kivistö & Pihlström 2020.

[20] This book is a rich collection of essays ranging from various historical explorations to theological and philosophical analyses of recognition phenomena in different historical and systematic contexts.

and love (focusing on general human dignity, specific identities, and unique individual personhood, respectively) as the main dimensions of recognition (ibid., pp. 1-2).

The editors continue to label recognition "a fundamental normative phenomenon" and to suggest that it "constitutes an adequate *response* to specific aspects of personhood" and may even play a crucial role "in the very *constitution* of general personhood, as well as more specific aspects of it" (ibid., p. 2). While the paradigmatic case of recognition is "a mutual granting of positive statuses between individual human persons", recognition extends to social groups as well as "normative entities quite generally" (ibid., p. 2). Fortunately, we are also reminded that although recognition is generally positive and a "vital human need", it has a "darker side" due to misrecognitions, power relations, and the need to "struggle" for recognition (cf. again Honneth 2005). Kahlos, Koskinen, and Palmén also take what we may call a "realistic" attitude to recognition by claiming that though it was Hegel whose work signals a turning point in the development of recognition theory, "the phenomena themselves were already present before their conceptual articulation by Hegel", because recognition is, indeed, a "basic human need" and presumably even constitutive of human persons and their identities (Kahlos et al. 2019, p. 4).[21]

Given the task of this paper, we again need to ask *how* "fundamental", exactly, recognition is as a normative phenomenon. One obvious question related to this general issue is how far the recognition theorist needs to go in the direction of realism. Would it be possible to maintain that recognition "phenomena" are, though "real", themselves something constituted (e.g., by further recognition acts)? They are themselves social phenomena, after all. This is a more general question regarding realism about the normative (as well as about historical social facts and institutions). In this context, however, a possibly more serious philosophical question can be formulated on the basis of the overview of normativity sketched above. No matter how "fundamental" recognition is

[21] Saarinen's (2016) historically detailed study also emphasizes that in theological and religious contexts recognition has been conceptualized in the history a long time before Hegel.

as a "normative phenomenon", it can be claimed that it is only *possible* within a context that is itself already richly normative. Perhaps the recognition theorist seeks to emphasize such irreducible normativity by suggesting that social reality is constituted by recognition acts, but my transcendental worry is that the very identity of those acts as recognition acts (rather than acts of some other kind) already presupposes a normative context.

Another issue the above-quoted comments raise is related to the strong emphasis on personhood among many recognition theorists. Does the world, we may ask, somehow divide itself up to, e.g., persons and non-persons already prior to recognition acts? Or do those acts (as is occasionally suggested) *constitute* persons (etc.) in a strong ontological sense?[22] But then how is it determined what kinds of things *can* through recognition acts be turned into persons? Some kind of "pre-recognition" must arguably have taken place for *relevant kinds of beings* to be even potentially recognizable as persons rather than something else. This, in turn, presupposes criteria of relevance that must, again, already be regarded as normative. Therefore, there just is no way to ground normativity in mere psychological acts of recognition, or contingent psychological acts of any kind.

Recognizing other human beings as persons, Koskinen (2019, p. 36) notes, involves acknowledging "their normative status as persons". Koskinen also refers to Robert Brandom's notion of "robust recognition", "the practical attitude of recognizing another as a simple recognizer", "as itself the kind of thing for which things can have a specifically normative significance" (ibid., p. 40). However, as recognition theorists like Koskinen of course clearly acknowledge, the normative form of life we share with other human beings may require (or even normatively obligate) us to recognize human beings for whom things do not, and cannot, have any normative significance because they lack the capacity to attribute such, or any, significance to anything. When Brandom and Koskinen

[22] Would our recognizing another as a person constitute their personhood also if we (or just I?) recognized animals, machines (artificial intelligence, robots), Martians, or the replicants familiar from the film *Blade Runner* as persons? Where would, or could, we draw the line?

characterize interpersonal recognition as an act of "[t]aking something to be subject to appraisals of its reasons, holding it rationally responsible" and thus of "treating it as some*one*: as one of *us* (rational beings)" (ibid., p. 42, quoting Brandom), the immediate issue that arises is how we should account for our recognizing human beings who are *not* persons, and not, except perhaps potentially but (tragically) not actually, among "us", such as the permanently ill or severely mentally dis-abled?

It is, I would like to argue, only within an always already normative context guided by something like (among others) the idea of *human dignity* – or some suitably general and irreducibly normative equivalent – that we can so much as *ask* whether, and how exactly, our various acts of recognition, or our failures to commit such acts, are appropriate or inappropriate, acknowledging or constituting relevant normative statuses. Dignity is transcendentally presupposed by any consideration of recognition vs. non- or misrecognition. It is, in short, only within a human form of life that is already thoroughly ethical and normative that we can discuss whether, and how, to recognize someone or something as something (and why). If this is the case, the human form of life in its normative dimensions *just cannot arise from (mere) recognition*. We must be human in order for us to be able to engage in recognition acts.[23]

In my view, these remain open issues; this paper admittedly raises more questions than it provides answers to. It is, in any event, unclear to what extent recognition is (or is even claimed by recognition theorists like Koskinen to be) a "fundamental" normative phenomenon in the sense that it could be taken to ground the normative order we live in, or the human form of life as such. I have suggested that contingent recognition acts are less fundamental than our already finding ourselves committed to and guided by normativity, be-

[23] "Let us be human", Wittgenstein once wrote (1980, p. 36), perhaps indicating that being human is already a task, something that normatively challenges *us* (as humans) only from within a human form of life – otherwise the encouragement would hardly make sense. (This phrase, like many Wittgensteinian ones, is thus arguably much more complex than its apparent simplicity might lead us to think.)

cause we need to so conceptualize our lives in order to be able to engage in any such acts in the first place. Therefore, normative statuses cannot be ultimately constituted by such recognition acts. But I also acknowledge the possibility that recognition theory might actually seek to express the kind of notion of dignity that I am invoking, because the affirmation of dignity as normatively foundational may itself be regarded as a recognition act.[24] If so, then recognition theory would already presuppose normativity more or less along the lines suggested in this paper – in which case the transcendental criticism of recognition would lose much of its relevance.

Be that as it may, recognition is arguably slightly less fundamental than we might be tempted to think, but we should be open to elaborations of recognition theory that render it closer to the transcendental requirements for "grounding" I have emphasized here. Furthermore, one might, in Wittgensteinian terms, also argue that recognition acts are always (for better or worse) somehow "reasoned" or "ratiocinated", while our being committed to normativity in general is, rather, based on "blind" rule-following, on our being "naturally" (though obviously not in the sense of reductive naturalism) engaged in the kind of practices within which our language-games find their homes.[25] A pragmatist (Peircean) version of this criticism would emphasize that particular recognition acts presuppose a wider context of habituality that is, again, already normatively structured.

In the remainder of this paper, I will try to further illuminate these issues by briefly returning to philosophical anthropology and then by taking a look at how "holistic pragmatism", a specific development of pragmatism, views the relation between the normative and the natural.

Dignity and the threat of refactualization

In contrast to any position that defends, say, human equality (understood as ethical and/or political or more generally normative) or basic human rights on the basis of merely natural and contingent facts about human beings, I would be will-

[24] Thanks are due to one of the referees for this very important point.

[25] "Language did not emerge from some kind of ratiocination", Wittgenstein (1969, §475) writes.

ing to suggest, drawing inspiration from Veikko Launis's (2018) recent work on human dignity, that the category of the human being, and the related category of human dignity, should be treated as more foundational than the category of the person, or the concept of "human rights".[26] I have argued above that our moral and generally normative reality cannot be transcendentally (regarding the conditions of its *possibility*) grounded in our acts of recognition even if such acts are constitutive of the normative and the social (and of personhood) in an ontological sense. In order for such acts themselves to be *possible*, we must, I have suggested, live in a thoroughly normative sphere in which we, for instance, evaluate any morally relevant acts and uses of language, including our recognizing behavior, in terms of our being *already* responsive to human dignity. This normative sphere is not reducible to contingent recognitions of personhood based on natural capacities, and it also invokes a notion of humanity wider than the category of the person, because we need to treat with dignity also those human beings who clearly lack the rational and other capacities of persons (e.g., deeply mentally disabled people). Our responsibility of treating others with dignity does not arise from our psychology or brain structure. It is, as Wittgensteinians might put it, *there* – "like our life".[27]

A worry that now rearises is whether our transcendental notion of human dignity, or any other normative notion we might use in a comparably fundamental (transcendental) normative role, is just a *cultural specificity* based on particular recognition acts we commit in our local cultural surroundings. Is it merely a local cultural practice, ultimately reducible to mere facts about what we in this specific culture do, to treat other human beings as equal?[28] This question brings us

---

[26] Launis's (2018) comprehensive work on human dignity is available only in Finnish. On the irreducible significance of the notion of the (other) human being in our lives, see, e.g., Gaita 2000.

[27] Cf. Wittgenstein 1969. In this sense, Sparti's (2005) way of speaking about the "responsibility" for acknowledging as more fundamental than acknowledging itself sounds somewhat problematic (and, despite his Wittgensteinian approach, curiously non-Wittgensteinian).

[28] This could be taken to be a problem analogous to the issue of epistemic and/or scientific norms of rationality being based on the contingent reasonings by scientists and other inquirers in specific historical contexts.

back to the issue concerning the relation between normativity and "human nature". Is there a kind of normativity already in place that enables us to, for example, ask the question whether it is our moral duty to avoid sexism or racism, or does our contingent recognizing the normative statuses of (say) women and people representing "other" ethnic backgrounds create any normativity there is in matters like this? We are back in basic issues of philosophical anthropology all over again.

Returning to Kannisto's (1984) scheme briefly introduced above, we may recall that cultural relativism is a natural development of culturalism, with the alarming tendency to "refactualize" the culture-specific normative order. Even if "we" in our culture do recognize women and non-white people, for instance, as fundamentally equal to white men, and if we set this recognition as a universal model to be carried over into other cultures as well, are we still only dealing with a local specificity that can ultimately be reduced to a mere contingent fact about how we behave and how we happen to think others ought to behave, too? How exactly should the relation between the natural and the cultural (or the contingently factual and the normative) be understood?

Pragmatism, as pointed out above, is an attempt to bridge the gap between the natural and the cultural, and therefore we should, before concluding the discussion, take a quick look at a promising pragmatist way of dealing with normativity.[29]

---

[29] Another relevant pragmatist approach *not* to be explored here would be Rorty's *ethnocentrism*, according to which we just have to "start from where we are", i.e., where we contingently find ourselves, and develop our "vocabularies" with the "ironic" awareness of the contingency of that starting point. We should not, I think, assess Rorty's pragmatism purely negatively; his emphasis on the historical contingency of our most fundamental normative frameworks is, I think, to be taken very seriously. I am, however, looking for a pragmatist account that would be reconcilable with a transcendental inquiry into normativity, and here Rorty seems to of little help – a form of pragmatism more responsive to the transcendental "vocabulary" is needed.

Holistic pragmatism and normativity

One suggestion for a way of developing a pragmatist philosophical anthropology entangling naturalism and normativity is Morton White's (1956, 2002) *holistic pragmatism*, which is basically an epistemological position but can be extended to a more general account of the "human form of life" (cf. Pihlström 2011, 2015). In a Quinean manner, White labels his pragmatism "holistic"; indeed, like his long-time friend and colleague W.V. Quine, he follows the anti-Cartesian and more generally anti-rationalist line of pragmatist thought (White 2002, 3-5), abandoning any "first philosophy". The distinctive character of White's position naturally emerges against the background of Quine's more extreme (and better known) views. While both Quine and White begin from a firm rejection of the analytic/synthetic distinction and from the holistic idea that our beliefs (or sentences we assent to) are not tested individually but "face the tribunal of experience" in corporate bodies, they draw quite different morals from this picture.

Whereas philosophy of science was, for Quine, "philosophy enough", White recommends that we extend holism from the philosophy of science to *philosophy of culture*, thematizing not only science but also other normative practices, such as religion, history, art, law, and morality (ibid., x-xi). This "cultural philosophy" covers philosophy of science as one of its subfields, but White insists that other cultural institutions require empirically informed philosophical scrutiny no less than science does (ibid., xiii). Holistic pragmatism maintains that "philosophy of art, of religion, of morality, or of other elements of culture is in great measure a discipline that is epistemically coordinate with philosophy of natural science" (ibid., p. 66). The idea that ethics, in particular, may be viewed as "empirical" if one includes feelings of moral obligation as well as empirical experiences in the "flux" of experience employed in the on-going critical testing of one's beliefs has been strongly present in White's writings from an early stage to the present (see White 1956, 1981, 2002). White is thus one of those philosophers who can be read as having defended a pragmatic form of moral and generally normative realism (cf. Pihlström 2005).

Quine (1953) took his famous holistic step by arguing that even logical truths are not immune to revision, because they are tested along with factual claims as components of larger conjunctions of statements (White 2002, p. 71). No general analytic/synthetic division can be drawn, as statements about, say, the synonymity of terms are ultimately empirical, describing the contingencies of factual language-use (ibid., pp. 71, 73). Despite this fundamental agreement with Quine, White argues against Quine that "observation sentences" (e.g., "That's a rabbit") and ethical sentences such as "That's outrageous" cannot be sharply separated from each other any more than analytic and synthetic statements can; their difference is a matter of degree, not a difference in kind (ibid., pp. 154-155, 160-163). Yet, ethical sentences at issue are genuinely normative:

> Avoiding the view that ethical sentences are synonymous with sociological or psychological sentences, and being impressed by the failure of reductive phenomenalism as well as the power of holism to bridge the traditional epistemic gap created by the distinction between the analytic and the synthetic, I propose a non-reductive version of holism in order to bridge the gap between the moral and the descriptive […]. (Ibid., p. 157.)

That is, descriptive statements and normative ethical principles form conjunctions that are tested holistically, just as Quine had argued that empirical and logico-mathematical beliefs in science are. Logic, science, and ethics form a unified whole, a holistic web without epistemic dichotomies. Moreover, as logical principles *may*, by Quinean lights, be given up in the face of sufficiently recalcitrant experience, descriptive statements *may* be denied in order to preserve a normative principle we do not want to give up (ibid., p. 159), although such situations may be rare. Hence, ethics is not immune to empirical evaluation, as feelings of obligation together with sensory observation link ethical sentences to the natural world. *Pace* Quine, ethics is, then, "anchored in experience" (ibid., p. 160). Furthermore, "feeling sentences" are fallible and can also be surrendered when a conjunction is tested (ibid., p. 166). Both ethics and science are, hence, corrigible yet cognitive elements of normatively structured human culture that in the end constitutes a holistic totality instead of

any compartmentalized group of distinct areas with definite boundaries. Knowledge and morals, as White himself put it many years ago, form a "seamless web" (White 1956, p. 287).

As an example of holistic pragmatism at work, consider the following argument:

(1) One can be morally responsible for one's actions (or have moral duties) only if one acts freely (i.e., is a genuine agent).

(2) One can act freely only if one possesses free will.

(3) Only individuals (can) have wills.

(4) Therefore, groups and collectives (e.g., business corporations) cannot have a will.

(5) Therefore, groups (etc.) cannot act freely and are not genuine agents.

(6) Therefore, groups (etc.) cannot be morally responsible.

(7) Therefore, there is no such thing as moral/social responsibility, nor any moral duties, attributed to business corporations.

Now, according to holistic pragmatism, if we find the normative conclusion ethically unacceptable, we may legitimately revise or reject one of the factual premises. Thus, if we find it ethically impossible to maintain, e.g., that business corporations cannot be responsible for their actions (or that we should not attribute such responsibility to them), we may revise our picture of what (their) agency (including freedom, the will, etc.), or agency generally, *is*. The revised picture of agency must then also be made compatible with the rest of our beliefs. Moreover, we must provide further reasons for the thoroughly normative ethical "impossibility" motivating this belief revision. There is, then, at least potentially, an endless process of *mutual holistic normative adjustment* of beliefs and evidence here – like in any empirical inquiry, yet extending to the fully normative sphere.

In comparison, consider this argument: (1) Racism is true. (2) If racism is true, then racial discrimination is justified. (3) Therefore, racial discrimination is justified. Now, to reject (3), we need not merely find purely theoretical or evidential reasons to reject (1). We may reject (1) because (3) cannot *work* as an element of our overall holistic system of belief within the

human form of life we find ourselves to be inhabiting. Our reason for rejecting a factual belief like (1) may be thoroughly normative, and holistic pragmatism makes sense of this.

I would be happy to construe these ideas somewhat more metaphysically as yielding the claim that there are, for us, no "value-neutral" facts at all (see Pihlström 2005, 2010), though I doubt that White himself ever intended them in such a metaphysical sense.[30] In any case, White's holism could be extended from the epistemic justification of different kinds of *statements* (sentences) or *beliefs* to whatever is the equivalent of such normative justification in the critical evaluation of entire cultural *practices* and normatively governed *institutions*. While remaining distinct from each other, such practices (e.g., science, politics, religion, art, and others) are dynamically interrelated and must therefore be evaluated holistically. A continuous critical (re)consideration of the normative structures that constitute our (thoroughly and irreducibly normative) form of life is precisely what holistic pragmatism calls for, and indeed makes sense of.

Another extension for holistic pragmatism is also needed because White's version is, arguably, *too thin*. Mere *feelings* of obligation are, again, just natural and contingent. Normative commitment to feeling-transcendent rational duty (in a quasi-Kantian sense) needs to be built into the holistic assessment of our normative-cum-factual belief systems. Moral emotions and even "mere" feelings do have a role to play here, but they cannot alone act as the epistemic ground for our moral commitments. And the same goes for more general normativity. This, however, would be a topic for another essay.

Conclusion: humanism

Holistic pragmatism is of course only one suggestion designed to meet our needs of defending the normative human form of life against reductively naturalizing (or "factualizing") tendencies, just as recognition theory has above been examined only as an example of an approach we might be tempted to employ in an attempt to ground normativity in

[30] White, like Putnam (2002), is strongly opposed to any metaphysical ("inflated") version of the fact-value entanglement. *This* might be seen as a remnant of logical positivism, too.

natural human capabilities and actions. In addition to "positive" suggestions seeking to articulate a pragmatist philosophical anthropology integrating (soft) naturalism and culturalism, it is at least equally important to engage in a "negative" critical examination of well-intended yet (in my view) insufficiently deeply normative proposals such as recognition theory. Let me now close with brief general remarks.

A defense of a normative order is – as my frequent references to the human "form of life" might also suggest – also a defense of a kind of *humanism*, even rather traditional Enlightenment humanism, with a reincarnation of the transcendental subject at its center, a subject self-reflectively examining its capacities and limits. This defense operates at a transcendental level: anti-, trans-, and post-humanism are all human beings' attempts to reflect on their relation to preestablished social and cultural hierarchies, non-human nature, animals, intelligent machines, etc. – to something non-human. A kind of *transcendental humanism* thus ultimately prevails, because any such criticisms of traditional humanism (just like any acts of recognition or arguments concerning whom to recognize, as what, and why) must inevitably take place within a space of reasons and thus within the human normative order. Only transcendental humanism makes empirical anti- or posthumanisms possible, analogously to the way in which for Kant it is only transcendental idealism that can make empirical realism possible.[31] Moreover, it is precisely on the basis of transcendental humanism that we can see the issue of normativity as inescapably – holistically – intertwined with the philosophical-anthropological question about what the human being is like.

I have argued for these conclusions by employing a negative philosophical method. In a more comprehensive discussion, it would be important to analyze critically not only the

---

[31] For these same humanistic reasons, I do not think the transcendental defense of dignity considered above extends to, say, non-human animals. But as our form of life changes, we might have to redefine what counts as "us", or even as humans. Even then, this would be a human change, and a human redefining process, in principle to be accounted for in terms of transcendental humanism.

horrible cases of the elimination or reduction of normativity to mere nature, such as the Nazis' reduction of the Jews to stateless and normless animals, to a kind of dehumanized indifference (cf. Levi 1988, Cavarero 2018), but also more "positive" reductionisms, such as the tendency to see the basis of morality in natural phenomena such as emotions.[32] The affirmation of human equality and dignity is, I have suggested, more fundamental than any contingent natural reactions (such as recognition) or any empirical evidence for or against contingent states of affairs. The commitment to valuing dignity *constitutes the normative sphere* within which (only) we can engage in the practice of discussing *anything* ethically or normatively at all.[33] Hence it cannot be defended (or criticized) by means of empirical evidence; it is more fundamental (like religious belief is for some Wittgensteinian philosophers of religion), albeit not in principle non-revisable or infallible. Thus, we could definitely end up in a dystopic world rejecting human dignity, though it would be difficult

[32] Analyses of the Nazi tendency to destroy the human (and thus normative) status of their victims, such as Cavarero's (2018), would benefit from an explicitly transcendental approach. For example, the very unforgiveability of the Nazi crimes may be seen as a transcendental insight into what the Nazis did: "Wherever the human is injured, human beings can neither forgive nor punish this radical offense to the human as such." (Ibid., p. 139.) It is self-evident that the Nazis had their own normative system, but a transcendental analysis may point out how deeply they were engaged in the dehumanizing project of destroying their victims' normative statuses – including the language in terms of which their human form of life had been meaningful to them. (For transcendental engagements with the problem of suffering, see Kivistö & Pihlström 2016.)

[33] This could be regarded as a reformulation of what I have elsewhere called "pragmatic moral realism" based on a transcendental argument (Pihlström 2005). Again, the recognition theorist could respond that making *this* claim is itself an act of recognition. There is no need to deny this, but one way of rephrasing my point is to suggest that the "always already" presupposed acknowledgment of normativity as a transcendentally pervasive feature of the human world may (when analyzed from an external perspective provided by, say, recognition theory) be realized as empirical (factual) acts of recognition, just as our transcendental self can be seen as identical to our empirical psychological self (i.e., not as an ontologically additional entity on top of the natural world). Cf. again Pihlström 2003, 2016.

for us to (now) include ourselves in that potential "we" that would have lost the transcendental framework of dignity inescapably characterizing our form of life. Moreover, holistic pragmatism reminds us that the boundary between the natural and the normative is itself constantly holistically tested and may historically change. Nothing, not even our normatively structured form of life conceived as a holistic totality, is beyond pragmatic critical transformation. This critical fallibilist spirit is itself inherent to humanism and to the normative framework that humanism defines.

Furthermore, it may also be acknowledged, at the meta-level, that transcendental inquiry into normativity is in an important sense *optional*; one *can* avoid it and engage in what I have called "naturalizing reduction" (or, less reductively, recognition theory) instead. In a sense the reductionist approach would, if my argument is on the right track, be somewhat like living in a transcendental illusion. Yet, such an illusion is visible *only from within*, i.e., only when we have made a transcendental turn and are committed to an "internal" analysis of normativity as constitutive of our lives. So whether the transcendental perspective *is* optional or not is a question receiving different answers depending on whether we have adopted that perspective or not. This reintroduces the relativism and refactualization issues all over again: our adopting the transcendental perspective in our inquiries is itself historically contingent, a local fact of the matter concerning our de facto processes of inquiry. And so it goes: the transcendental inquirer cannot avoid working within a kind of endlessly reflexive spiral.

A final note is needed. The transcendental problem concerning the very possibility of normativity is, we should admit, a philosophical mystery deep enough to make it understandable (albeit not for that reason justifiable) that some of us think it cannot be solved without reference to something *transcendent*.[34] However, the transcendental humanist maintains that even by making such a move we cannot get rid of our inescapably human starting point. Even

[34] Taylor (1989), for example, ultimately places his account of "strong evaluation" in a theistic context; cf. Pihlström 2011 for some critical remarks on the relation between the transcendental and the transcendent.

theism would not liberate us from the *burden* of humanism and the puzzlement about normativity. Normativity is an enigma *for us*. In philosophical-anthropological terms, a culturalist (humanist) view of the irreducible normativity of the human world is in a constant danger of collapsing into either cosmic transcendent teleology (classical or Christian) or refactualizing naturalism and/or cultural relativism, or perhaps the individual contingency of existentialism.[35] Transcendental humanism is needed at the meta-level to guide our search for plausible accounts defined by these open issues, and especially to guard us against too easy solutions.[36]

University of Helsinki

References

Arendt, Hannah (1958), *The Origins of Totalitarianism*, University of Chicago Press, Chicago.

Bernstein, Richard M. (2018), *Why Read Hannah Arendt Now*, Polity, Cambridge.

Cavarero, Adriana (2018), "'Destroy Your Sight with a New Gorgon': Mass Atrocity and the Phenomenology of Horror", in T. Brudholm & J. Lang (eds.), *Emotions and Mass Atrocity: Philosophical and Theoretical Explorations*, Cambridge University Press, Cambridge, 123-141.

Dewey, John (1929), *Experience and Nature*, Open Court, La Salle, IL, 1986.

Honenberger, Philip (ed.) (2016), *Naturalism and Philosophical Anthropology: Nature, Life, and the Human between Empirical and Transcendental Perspectives*, Palgrave Macmillan, Basingstoke.

Honneth, Axel (2005), *Kampf um Anerkennung*, 2nd ed., Suhrkamp, Frankfurt, 1st ed. 1992.

James, William (1907), *Pragmatism: A New Name for Some Old Ways of Thinking*, eds. Frederick H. Burkhardt, Fredson Bowers, & Ignas K. Skrupskelis, Harvard University Press, Cambridge, MA and London, 1975.

Kahlos, Maijastina, Koskinen, Heikki J., & Palmén, Ritva (eds.) (2019), *Recognition and Religion*, Routledge, London and New York.

---

[35] I examine these worries in more detail in Pihlström 2016.

[36] I am grateful to two anonymous reviewers as well as many of the participants of the I.I.P. Entretiens at the University of Helsinki (August, 2019) for highly valuable critical comments on this paper.

Kannisto, Heikki (1984), "Filosofisen antropologian mahdollisuudesta" [On the Possibility of Philosophical Anthropology], *Ajatus* 41.

Kant, Immanuel (1781/1787), *Kritik der reinen Vernunft*, ed. Raymund Schmidt, Felix Meiner, Hamburg 1990.

Kivistö, Sari & Pihlström, Sami (2016), *Kantian Antitheodicy: Philosophical and Literary Varieties*, Palgrave Macmillan, Basingstoke.

Korsgaard, Christine (1996), *The Sources of Normativity*, Harvard University Press, Cambridge, MA.

Koskinen, Heikki J. (2017), "Mediated Recognition and the Categorial Stance", *Journal of Social Ontology* 3, pp. 67-87.

Koskinen, Heikki J. (2019), "Mediated Recognition: Suggestions towards an Articulation", in Kahlos et al. 2019, pp. 34-50.

Launis, Veikko (2018), *Ihmisarvo* [Human Dignity], Vastapaino, Tampere.

Levi, Primo (1988), *The Drowned and the Saved* (1986), trans. Raymund Rosenthal, Michael Joseph, London.

McDowell, John (1996), *Mind and World,* 2nd ed. (1st ed. 1994). Harvard University Press, Cambridge, MA and London.

Pihlström, Sami (2003), *Naturalizing the Transcendental: A Pragmatic View*, Humanity Books, Amherst, NY.

Pihlström, Sami (2005), *Pragmatic Moral Realism: A Transcendental Defense*. Rodopi, Amsterdam.

Pihlström, Sami (2010), "Toward a Pragmatically Naturalist Metaphysics of the Fact-Value Entanglement: Emergence or Continuity?", *Journal of Philosophical Research* 35, 323-352.

Pihlström, Sami (2011), "Morton White's Philosophy of Culture", *Human Affairs* 21, 140-156.

Pihlström, Sami (2014), *Taking Evil Seriously*, Palgrave Macmillan, Basingstoke.

Pihlström, Sami (ed.) (2015), *The Bloomsbury Companion to Pragmatism*, Bloomsbury, New York and London.

Pihlström, Sami (2016), *Death and Finitude: Toward a Pragmatic Transcendental Anthropology of Human Limits and Mortality*, Lexington, Lanham, MD.

Pihlström, Sami (2020), *Pragmatic Realism, Religious Truth, and Antitheodicy: On Viewing the World by Acknowledging the Other*. Helsinki University Press, Helsinki.

Plato, *The Republic*.

Putnam, Hilary (1981). *Reason, Truth and History*, Cambridge University Press, Cambridge.

Putnam, Hilary (1983), *Realism and Reason*, Cambridge University Press, Cambridge.

Putnam, Hilary (1999), *The Threefold Cord*, Columbia University Press, New York.
Quine, W.V. (1953), *From a Logical Point of View*, rev. ed., Harvard University Press, Cambridge, MA and London, 1980.
Rorty, Richard (1989), *Contingency, Irony and Solidarity*, Cambridge University Press, Cambridge.
Saarinen, Risto (2016), *Recognition and Religion: A Historical and Systematic Study*, Oxford University Press, Oxford.
Snyder, Timothy (2010), *Bloodlands: Europe between Hitler and Stalin*, Vintage London.
Snyder, Timothy (2015), *Black Earth: The Holocaust as History and Warning*, Vintage, London.
Sparti, Davide (2005), "Let Us Be Human: Primo Levi and Ludwig Wittgenstein", *Philosophy and Literature* 29, pp. 444–459.
Taylor, Charles (1989), *Sources of the Self: The Making of Modern Identity*, Cambridge University Press, Cambridge, 1994.
Taylor, Charles (1992), "The Politics of Recognition", in Taylor, *Philosophical Arguments*, Harvard University Press, Cambridge, MA and London, 1995.
White, Morton (1956), *Toward Reunion in Philosophy*, Harvard University Press, Cambridge, MA.
White, Morton (1981), *What Is and What Ought to Be Done*, Harvard University Press, Cambridge, MA.
White, Morton (2002), *A Philosophy of Culture: The Scope of Holistic Pragmatism*, Princeton University Press, Princeton, NJ.
Wittgenstein, Ludwig (1953), *Philosophical Investigations,* trans. G.E.M. Anscombe, Basil Blackwell, Oxford, 1958.
Wittgenstein, Ludwig (1969), *On Certainty*, trans. G.E.M. Anscombe & D. Paul, Basil Blackwell, Oxford.
Wittgenstein, Ludwig (1980), *Culture and Value*, trans. Peter Winch, Basil Blackwell, Oxford.