Norms in Social Interaction : Semantic, Epistemic, and Dynamic

Lo Presti, Patrizio

2015

Norms in Social Interaction:
Semantic, Epistemic, and Dynamic

Patrizio Lo Presti

LUND
UNIVERSITY

Faculty of Humanities | Department of Philosophy
ISBN 978-91-87833-66-3 (Print)
ISBN 978-91-87833-67-0 (Pdf)

Printed in Sweden by Media-Tryck, Lund University
Lund 2015


To Amanda,
Inger, Angelo, Tina, Nicola and Micke,
family and friends.

[T]he intelligent reasoner [...] reasons with a correct method, but without considering the prescriptions of a methodology. The rules that he observes have become his way of thinking, when he is taking care [...]

The boxer, the surgeon, the poet and the salesman apply their special criteria in the performance of their special tasks, for they are trying to get things right; and they are appraised as clever, skillful, inspired or shrewd not for the ways in which they consider, if they consider at all, prescriptions for conducting their special performances, but for the ways in which they conduct those performances themselves. Whether or not the boxer plans his manoeuvres before executing them, his cleverness at boxing is decided in the light of how he fights. If he is a Hamlet of the ring, he will be condemned as an inferior fighter, though perhaps a brilliant theorist or critic. Cleverness at fighting is exhibited in the giving and parrying of blows, not in the acceptance or rejection of propositions about blows, just as ability at reasoning is exhibited in the construction of valid arguments and the detection of fallacies, not in the avowal of logicians' formulae. Nor does the surgeon's skill function in his tongue uttering medical truths but only in his hands making the correct movements.

Gilbert Ryle, *The Concept of Mind* (1949/2009), p. 36.

Contents

	Acknowledgements	9
	List of Papers	13
1	Introduction	15
1.1	Background on the Variety of Norms	15
	1.1.a Social Norms	16
	1.1.b Semantic Norms	19
	1.1.c Epistemic Norms	22
1.2	Strategy and Work Plan	24
1.3	Norms and Analysis	26
2	Theory, Definition, and Method	31
2.1	Background on Models: Cognitive and Non-Cognitive	31
2.2	The Non-Cognitive Model and Definitions	32
	2.2.a Embodiment	33
	2.2.b Situatedness	37
	2.2.c Ecology	41
	2.2.d Enaction and Dynamics	44
	2.2.e Summary	47
2.3.	The Cognitive Model	48
2.4	Method and Hypothesis	52

3 Summary of the Papers 59
Paper 1 59
Paper 2 60
Paper 3 62
Paper 4 63
Paper 5 65
Paper 6 66

4 Conclusions 69

References 73

Papers 1-6 79
Paper 1 79
Paper 2 105
Paper 3 121
Paper 4 135
Paper 5 169
Paper 6 201

Acknowledgements

Work on this dissertation was funded by the Swedish Research Council (grant no. 429-2010-7181) and the European Science Foundation as part of the EUROCORES project "Understanding and Misunderstanding: Cognition, Communication and Culture" (EuroUnderstanding). This dissertation is a result of the "Understanding Rules: Cognitive and Non-Cognitive Models of Social Cognition" project, which was one of three in the "Understanding the Normative Dimensions of Human Conduct: Conceptual and Developmental Issues" (NormCon) project. The project's principal investigator in Lund was Professor Ingar Brinck.

Funding has also been received from the Knut and Alice Wallenberg's trust (grant no. RFh2013-0465), Spouses Engineer Lars Henrik Fornander's trust (grant no. FO2013-0004), Erik and Gurli Hultengren's trust for philosophical research (grant no. Huh-2012-0012), and travelling and research grants from the Faculty of Humanities at Lund University. These funds allowed exchanges with the philosophy and psychology departments at Salzburg University, the philosophy departments at the University of Warwick and Stockholm University, as well as attendance to several conferences: the ESF conferences in Malmö 2011 and Lisbon 2014, the ESPP conferences in London 2012 and Noto 2014, The Minds in Common conferences in Aarhus 2012 and Paris 2013, the Phenomenology of Sociality conference in Dublin 2013, the Swedish Congress of Philosophy in Stockholm 2013, the ECAP conference in Bucharest 2014, and the NormCon meetings in Salzburg 2012, Copenhagen 2013, and Lund 2015.

I want to express my gratitude to the audiences at these conferences and to the funders for the profitable exchanges of ideas they sponsored.

Ingar Brinck, above all, should be lauded for her patient, sharp, kind, challenging, inspiring, day-and-night, every-day supervision. It is not without a feeling of dread that I think about the half-baked and hardly digestible manuscripts she has had to read the last four years. If passing from undergraduate to graduate studies and dissertation is a challenge, supervising the processes now seems to me perhaps the greater challenge. Ingar's and mine several discussions on normativity, rules, and social cognition have profoundly influenced this dissertation.

I am also grateful to my assistant supervisor Björn Peterson. Björn's expertise on social ontology and collective agency was crucial for several of the papers reprinted in this volume. The calm and carefulness with which Björn and I have discussed these topics has not only helped me develop some of the arguments of the dissertation but has also become something of a role model for philosophical discussions.

Special thanks also goes out to the other members of the NormCon project. To Johannes Brandl for welcoming me in Salzburg, for our discussions on rule-following, for his concern with the progression of my work, and for his accepting to be external discussant of this dissertation at the final seminar in June 2015. To Frank Esken for inviting me to give a presentation in Osnabrück. To Josef Perner, Hannes Rakozcy, Beate Priewasser and Eva Rafetseder for the discussions on early-in-development understanding of rules. To Johannes Roessler for our meetings in Warwick. To Dan Zahavi and Glenda Satne for arranging the NormCon workshop at the Centre for Subjectivity Research in Copenhagen. To Åsa Wikforss, not a member of NormCon, for her meticulous readings and comments on my work on epistemic and semantic normativity. To Pascal Engel, also not member of NormCon, for his comments on the same topics.

Editors and reviewers at the journals in which parts of this dissertation are published should be recognized for their work and for their permissions to reprint.

The CogCom Lab, led by Ingar, has been of much help. Thanks the members of this group: Susanna Bernstrup, Andreas Falck, Åsa Harvard, Elaine Madsen, and Thord Svensson. The participants and colleagues at

the Higher Seminars of the Theoretical Philosophy division, led by Erik J. Olsson, should all receive a thank you: Staffan Angere, Sebastian Enqvist, Emmanuel Genot, Jens Ulrik Hansen, Bengt Hansson, Tobias Hansson Wahlberg, Justine Jacot, Ingvar Johansson, Martin Jönsson, Rasmus Kraemmer Rendsvig, George Masterton, Johannes Persson, Carlo Proietti, Paula Quinon, Stefan Schubert, Jeroen Smid, and Frank Zenker. In addition I am grateful to the members of the Metaphysics and Collectivity group, in addition to those already mentioned: Gunnar Björnsson, Olle Blomberg, Johan Brännmark, Anna-Sofia Maurin, Susanna Salmijärvi, and Andras Szigeti.

For kind and swift help with tiresome technical and administrative work I would like to thank Kim Andersen, Agneta Ahlberg, Anna Cagnan Enhörning, Richard Johansson, Tomas Persson, Jesper Olsson at the Humanities office, Ylva von Gerber, Anna Östberg, and the Lund University IT-support unit. Thank you also Åsa Burman for the Finish On Time course. I am grateful for the friendly help of Jan Hartman.

Last but far from least I would like to thank my doctoral student colleagues Asger Kirkeby-Hinrup and Oscar Ralsmark for our lively discussions in the philosophy of mind reading group as well as for helping me with comments on several occasions. Thank you Frits Gåvertsson for all those pauses outside the Department of Philosophy at Kungshuset and LUX. And thank you Martin Viktorelius, friend since undergraduate studies and now doctoral student in cognitive science at Chalmers University of Technology. None of this would be possible without the last eight years of friendship and off the hook, open-minded, at the bar, at the dinner table, on the train, and on vacation discussions of everything in philosophy, and what it is that we are really doing when doing philosophy.

List of Papers

1. Lo Presti, P. (2013). Situating Norms and Jointness of Social Interaction.
   Published in: *The Journal of Natural and Social Philosophy*, vol. 9, no. 1, pp. 225–248.

2. Lo Presti, P. (2013). Social Ontology and Social Cognition.
   Published in: *Abstracta*, vol. 7, no. 1, pp. 5–17.

3. Lo Presti, P. (2015). Rule-Following, Meaning Constitution, and Enaction.
   Published in: *Human Affairs*, vol. 25, pp. 110–120.

4. Lo Presti, P. (In Press). An Ecological Approach to Normativity.
   In press in: *Adaptive Behavior*.

5. Lo Presti, P. (Submitted). Speaking About the Normativity of Meaning.
   Submitted to: *Philosophical Studies*.

6. Lo Presti, P. (2014). Moore's Paradox and Epistemic Norms.
   Published in: *Logos & Episteme*, vol. 5, no. 4, pp. 445–464.

1. Introduction

This dissertation asks how people learn, understand, and act according to norms. A norm may be broadly construed as what it is correct, right or appropriate to do. This dissertation treats the problem how people learn, understand, and act as is correct. Thus understood, norms apply in many domains of conduct, in different context, and at many levels. Norms may apply to all from how to meet another's gaze to how to do arithmetic in different contexts, such as those of a dinner with friends to that of teaching a primary school class how to do addition. And one may understand that it is inappropriate to meet another's gaze in a particular manner, at the level of being able to say that it is inappropriate, while, at the level of actually avoiding doing something inappropriate, fail to do so in an actual case because one is dumbfound.

In this chapter I give a background to the variety of meanings that the term 'norm' has been given and how I go about investigating how norms are learned, understood, and acted in accordance to. In the next chapter I present the theoretical background that this dissertation investigates and contributes to, define central terms, and discuss method (chapter 2). I then summarize the papers and conclusions reached (chapters 3 and 4).

1.1 Background on the Variety of Norms

There are many different meanings of 'norm'. The different meanings correspond to different domains of human conduct in which the term is thought to have reference, such as that of face-to-face social interaction,

language use, belief and epistemic agency. When we look closer at these we will see that their variety suggests that there is not one homogeneous sense of the term 'norm', but many—perhaps as many as there are domains of practices with which norms are associated.

In this dissertation the normativity of social interaction, meaning, and belief is investigated. It is asked, first, if norms apply in the given domain and, if so, how we should understand in what sense norms apply: necessarily or contingently. This section gives a short background to the subject matter in what domains of human conduct norms have been thought to apply and in what sense they have been taken to do so.

1.1.a Social Norms

When it is said, e.g., that there is a correct standing distance in conversation, this might be taken to mean that there is a norm how close to stand to one's interlocutor in conversation. What this means is that if one stands closer or further away the other is in position to correct one, to provide positive or negative feedback with respect to how close one stands. This we might call a *social* norm, pertaining, as it does, to social interaction.

Other social norms might involve those for: standing in line, leaving one's seat on the bus for the elderly, littering, spying on others, biking through the park, dress-codes, etiquette, table-manners, eye contact, etc. Some social norms might be explicitly regulated, as is the case, for instance, when a sign is put up that says, "Do not litter!" Norms may also be implicit, as when people regard biking through the park to be incorrect although there is no explicit agreement that it is incorrect.

Social norms might or might not be equivalent to explicit regulations and laws. For example, a law that prohibits begging might be regarded as incorrect in a community. People might not correct or sanction people for begging, and perhaps they would regard sanctioning beggars as incorrect. The sign at the entrance of the park that says "No biking allowed!" might be ignored and almost no one regard biking through the park as incorrect. Indeed, people might be baffled if they were corrected for biking through the park, because there is an implicit agreement that the explicit rule does not deserve compliance.

Social norms might be understood as *dynamic* (Brinck 2015). This means that the norm is in constant change as a consequence of what people do in the context of participation in the practice with which the norm is associated. Social norms might also be understood as dynamic in the sense that they tend to produce changes in the pursuit of activities with which they are associated. A norm for standing distance in conversation, e.g., might change during the day because of circumstantial noise, such as that of traffic. Late at night it is taken to be correct to stand at a certain distance, but if it is difficult to hear, because of rush-hour traffic, to move closer is taken to be correct. The same can be said for other social norms. For example, a norm not to litter might change over time as a consequence of people's littering-behaviour. People living in a littered environment might perceive this as the absence of a no-littering norm. But if some are observed to litter less this might be taken by others to be the beginning of a change of the norm. Others adapt to the behaviour, with the consequence that a no littering-norm is adopted.

The sense in which norms are 'dynamic' that I am after here is this: they are in all circumstances of interaction in practices with which they are associated in constant, non-linear, change. That change is non-linear means, in the terminology of dynamic-systems theory (Thompson 2007, pp. 38-43), that its progression in time is non-predictable from any given specification of initial conditions at a discrete instant. To say that social norms are dynamic in this sense is to say that the problem to specify at an instant, taken as initial condition, what will be the norm at some future instant, taken as the output state from progression of social interaction in the relevant practice, has no "analytical solution". An analytical solution to the problem of specifying change is one where all future states of a system can be known given a specification of its initial conditions. Analytical solutions are possible only for linear change over discrete instants where all initial conditions are given. If social norms are dynamic it follows that no individual's understanding of what is correct to do can consist in her inferring, from a specification of an instant of the practice with which the norm is associated, what is correct at some other instant, because this requires that the progression of interaction between the two instants is linear; i.e., that no change not given by a specification of the initial conditions is introduced as interaction progresses.

Changes in social norms may come about for a variety of reasons. For example, the passing of a law defining littering as punishable might lead people to litter less and regard it as incorrect to litter. But it also might not, because people continue to regard littering as permissible despite the law. Perhaps they even regard it as correct to violate the law. Thus the dynamics of norm-change might be understood as the working of both top-down and bottom-up influences; influenced from above by authorities such as law-makers or the state who impose rules and edicts, and from below by what people actually do and how they implicitly agree (or disagree) in behaviour.

A central question that this dissertation asks is whether norms in general are to be understood as dynamic and, if so, how a theory of norm-dynamics may be formulated.

Social norms might quite straightforwardly be construed not only as dynamic but also as *community-relative*. Norms for littering might be quite different between, e.g., Swedes and Italians, but they might also be quite different between subgroups within a wider, common community. Even within subgroups different people might regard different behaviours as correct. Insofar we are concerned with social norms, however, they apply minimally to a dyad of two people interacting. Below the dyad we have private opinions, and these do not constitute norms because one is not in a position to correct someone for failing to act in accordance with a social standard for correctness if that standard turns out just to be what one person *thinks* is correct (Wittgenstein 1953/1958, §202).

On the other hand, insofar we are concerned with social norms there seems to be no wider context of application than *all* communities (Wright 1980). There is no social standard outside the widest social context constituted by all communities. A social norm that applies in all communities cannot from some further, extra-social point of view, be regarded as correct or incorrect. It is in this sense universal. Indeed, it appears oxymoronic to say that a social norm is valid according to some non-social authority; according to some individual's private opinion or according to some point of view external to all communities. But a social norm can always be counteracted or confirmed from within, because of the dynamics of social interaction in various communities. Thus there is the highest macro-level—the community of all communities—from which no *higher* standard can be appealed to, and there is the lowest micro-level—the dyad—from which no

lower standard can be appealed to. As Margaret Gilbert (2003) puts it, we might think of two individuals in interaction as the "social atom"—the smallest context to which social norms can apply. Likewise we may think of all communities together as the "social universe"—the widest context to which social norms can apply. In between there are intermediaries; smaller dyadic, triadic, ..., *n*-adic social contexts, interactions within which dynamically change or stabilize norms in a dialectical process in which implicit agreement is established, changed, and abandoned and on which explicit regulation can be imposed.

This dissertation contributes with an investigation on the social nature of norms and social understanding; how these are realized in behavioural patterns in social interaction or in collective (explicit or implicit) agreement. Coupled with the investigation into the dynamic nature of norms, this dissertation contributes to a socio-dynamic understanding of norms.

When I discuss the other varieties of norms below it might be supposed that they are social and dynamic, or that they are not. For example, as will be seen, some authors have insisted that the alleged epistemic norms for belief apply irrespective of whether anyone accepts or agrees to them and that they do not change. The same can be said about the alleged norms for meaning and rationality.

This dissertation contributes primarily to a theory of how people learn, understand, and manage to act according to social norms.

1.1.b Semantic Norms

A common position among philosophers is that what distinguishes noises and marks that *have meaning* from those that do not is that the former are subject to correctness conditions while the latter are not. This means that the noises I make and the ink-marks on this paper have meaning only if one can assess whether they are correctly or incorrectly used relative to some condition, e.g., that they are grammatically well-formed, that what I mean by them correspond to what they are taken to mean in our community, and that they are used in a manner appropriate in our community.

Now, correctness, it seems, is a *normative* term; it implies that there are certain things one *ought (not)* or *should (not)* do (Gibbard 2014). So meaning must be normative if it entails correctness conditions; that a word

has a certain meaning means that one should use it according to certain norms. This is called meaning-normativism and has its modern roots in Saul Kripke (1982), who argued (or argued that Wittgenstein argued) that the relation between meaning and use must not be how we *do* or *will* use language according to some pattern but how we *should* use it.

On a competing approach to meaning and normativity correctness is not unambiguously a normative term (Hattiangadi 2007). It can have descriptive purport as well. For example, when it is said that 'dog' means *dog* implies that 'dog' can be used correctly (to refer to dogs) or incorrectly (to refer to non-dogs), then this only means that it is *true*, relative to those correctness conditions, that 'dog' means *dog*. It does not follow from this, or so it is argued by anti-normativists, that there is any particular manner one should or should not use 'dog'; words refer and to use words correctly simply means that it is true that one uses them according to a certain standard. Hence, it seems, 'correctness' has *descriptive* meaning, and norms must be *imposed*, e.g., socially, for meaning to be normative (Glüer and Wikforss 2009).

Anti-normativists agree that norms can be imposed on language use; i.e., it *can* be normative but it is not a conceptual truth that it is. For example, it might be a social norm that one should not lie, not assert what one believes to be false, to be conspicuous, and so on (Grice 1989). But these are not norms that can be directly derived from meaning alone. They must be supported by social agreement on correct use.

Hence to say that meaning entails correctness, as many philosophers do, is not, according to anti-normativists, to say that meaning is *normative*, because correctness might be a non-normative term. At least, it is ambiguous whether correctness is normative and therefore not sufficient for meaning to entail normativity that it entails correctness.

It can be seen here that there is a difference between saying, on the one hand, that meaning is normative in the sense that *semantics* is normative and, on the other hand, saying that *pragmatics* is normative. Thus it might be argued that the semantics of a language does not entail norms, because for norms to apply the language must have a socially agreed on pragmatics; that is, an agreement in use.

Given this difference between norms for semantics and norms for pragmatics several positions are opened up. For example, one might argue

that while meaning *alone* and in abstraction form *use* is not normative, the latter is normative. And if one accepts that meaning is constituted in use, then the claim that meaning is non-normative, while use is, must be understood as an abstraction from language of what is constitutive of it; namely, norms. On the other hand, one might reject that meaning consists in use, in which case one might reject that meaning is normative even if one accepts that use is.

This dissertation provides an investigation of the hypothesis that meaning is normative by means of analysis of the arguments presented on both sides of this debate. The hypothesis, more specifically, is that meaning is indeed normative in the sense that noise or marks have meaning if and only if users (speakers) can be relied on as committing to certain patterns in use. And so someone cannot be understood as a speaker if she does not exhibit sensitivity to commitments that expressions are implicitly undertakings of. On this view, that something has meaning entails that it implicitly commits and entitles to certain use. But that something has meaning does not entail that there are certain things one *ought (not)* or *should (not)* say *given* the meaning of the words one would then use. Given its meaning, one might do what one wants with language insofar as one can be recognized as using language (of course, one may fail to be so recognized as well—it is not the case that one *ought* or *should* make oneself recognizable as using a language). That is to say, norms for pragmatics are social and contingent and one may violate them and still be recognizable as a speaker. But what makes for this freedom in use are certain normative constraints, namely that one commits to, and can be relied on by others as committed to, certain patterns in use. If the hypothesis is validated, then it might be said: the fact that this text means something means that it can be read as committing the author to certain claims, and also as entitling the author to certain claims (even such the author does not recognize), and entitle the reader to criticise and correct those claims or to commit to them because the reader becomes convinced of what the author says. It does not mean that the author ought to or should write certain things. Insofar the reader can tell that the author is hereby committing and being entitled to certain claims, that is all it takes for this text to have meaning.

1.1.c Epistemic Norms

Saying that a belief is correct only if it is true suggests that there is a truth-norm governing belief. In the same manner, saying that it is correct to believe something only if one has evidence for it might mean that belief is governed by an evidence-norm. Then there might be a knowledge-norm for belief; namely, that it is correct to believe only what one knows.

These are all epistemic norms because they would apply to the formation, entertaining, and revising of one's epistemic position.

The relation between correctness and epistemic positions to which correctness applies might in turn be construed in many different ways. For example, it might be argued that it is a conceptual truth (Shah and Velleman 2005) that belief is normative. In this sense, assuming that the relevant epistemic norm is that of truth, to consciously believe, e.g., that it is raining while also consciously conceding that it is not true that it is raining is something one *cannot* really do (Engel 2013). The situation is such that we simply cannot understand the person as believing that it is raining and conceding that it is not true that it is raining. Moreover, the individual herself must in this case have misunderstood what it *means* to believe.

On the conceptual claim a norm applies quite regardless of whether anyone thinks or accepts that it does; hence it is not a matter of social agreement, as argued by Pascal Engel (2001). For example, if we in our community should come to accept that it is correct to believe what one also consciously and simultaneously concedes to be false, it is not the case, on this approach, that it is then correct to have such beliefs. And that it is correct to believe only what is true, if that is the norm claimed to apply by necessity, then it is not a fact that we take it to be correct to believe only what is true that makes all and only true beliefs correct. This would be the case regardless of what we take to be correct.

In contrast, the relation between correctness and epistemic positions might be construed as *(socially) contingent*. This means that the relevant norm, e.g., the truth-norm, applies to belief because we take it to apply. Thus, for instance, to believe what one lacks evidence for might be regarded as incorrect in a community because people in that community take belief to be something that should be supported by evidence. It is not that the norm applies as matter of *conceptual truth* in the previous sense,

because the norm might change or be abandoned over time as people's attitudes about what is correct change. So it is not the case that one *cannot* really believe something for which one lacks evidence. Rather, if one does so, others might demand that one correct one's epistemic position and, if one does not do so, they may simply ignore that position as insincere or as something that should not be taken into account.

On the latter, socially contingent construal of the application of norms to epistemic positions the relevant norms are social in the sense that there *is no norm for belief* if no one, minimally a dyad, thinks or accepts that there is and that it applies in their situation. This does not mean that the norm is not objective. It might be objectively true that one "ought to believe the fairly obvious consequences" of what one believes (Jackson 1999, p. 421) because this is agreed in a community (Searle 1995), even if the agreement can change and thus the norm be abandoned. So, to say that norms consist in social agreement does not imply that they are not objective.

Thus we can see clearly that the mere claim that belief is governed by norms opens up for a wide variety of positions with respect to, first, *what* the norm is supposed to be (e.g., a truth-norm or an evidence-norm) and, secondly, with respect to *how* the norm is supposed to apply (necessarily or contingent on social attitudes of taking the norm to apply).

This dissertation will contribute with an assessment of arguments that belief is normative in two specific senses in one context in particular; namely, the context of a certain class of paradoxical beliefs that have become known as Moore-paradoxical (Moore 1942). These are described as follows: on one version I believe that *P* and I believe that I do not believe that *P*; on a second version I believe that *P* and I believe that I believe that *not-P*. It will be claimed that arguments that such beliefs are paradoxical because the believer violates norms that necessarily apply to all beliefs fail.

In this section I have only wanted to show how rich in meaning the terms 'norm' and 'normative' turns out to be once we investigate the many domains in which norms might be thought to apply. And I have only mentioned a few of these.[1]

This introduction to the subject matter of normativity in various domains of human conduct should prepare the reader for the arguments and theses relevant to this dissertation. It also is meant to give a back-

ground on what philosophers have had to say about the role of norms in human conduct in general since the second half of the twentieth century.

1.2 Strategy and Work Plan

The above is intended to show just how wide the field of the meaning of 'norm' is. In this section I want to discuss what implications this has for the formulation of a general theory of norms, covering the whole field.

An obvious problem that has glared up on me from the vast literature on norms, as they might apply to all from a handshake to the solution of complex equations, is that the meaning of 'norm' is overwhelmingly heterogeneous. For this reason it has proved difficult to formulate a simple definition, in terms of necessary and sufficient conditions, stating: "*x* is a norm *if and only if* ..." From the point of view of analytic philosophy this is a nuisance. "What do you mean by 'norm'?" it will be asked. And, after a pause it might be continued "Well, if you don't have an answer then what in the world are you speaking about?" I would like to point out two things here.

First, we might have an ordinary, everyday understanding of a term. Even if we are not in a position to state what the necessary and sufficient conditions are for something to be a norm, we may still have a notion of it that is sufficiently clear for discussing and investigating phenomena in which norms are relevant. I will say more about this below.

Secondly, it might just be the case that because 'norm' applies in such a wide variety of domains of conduct it must be understood somewhat differently in each. In saying this I am not in bad company. As George H. von Wright writes in the opening to his *Norm and Action* (1963, p. 1):

> Since the field of meaning of 'norm' is not only heterogeneous 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.

From this observation, accurate or not, von Wright goes on to distinguish a host of different domains in which the term 'norm' might apply, much as I have done here.

I have chosen not to try to formulate a grand, General Theory of Norms in this dissertation. For the purpose of engaging, as I do in the papers that follow, in debates about what role norms plays in several different domains (social interaction, meaning, and belief), there simply seems to be no *one* meaning of 'norm' that one so much as could have opted for throughout. And yet, it seems, it is perfectly possible to go on and discuss what relevance norms have in these different domains, insofar at least that one does not simply presume that the meaning of the term is easily transferrable *salva significatione* across the whole field. For this reason I have chosen to favour specificity within each domain investigated instead of generality across the whole field, so as not to be the proverbial bull in a china shop.

I am afraid, therefore, that the reader who expects a crisp and general definition of 'norm' is likely to be disappointed. Nevertheless, and despite this disparaging situation, I maintain that the fact that the field of meaning of norm is varied is not sufficient for saying that our use of the term is a confused jumble. This is so because we might make out a *core* sense, even if the boundaries are somewhat vague.

Above I stated as the first reason not to be discouraged by the absence of a general analysis of norm that we might still work with an everyday notion sufficiently clear for the investigation that follows. Let me say something about this everyday notion here.

Unless what is disputed is the very claim that norms tell us what it is correct to do or standards allowing assessments of the appropriateness of a deed (as it is in paper 5), we might say that *that* is precisely what norms, in an everyday sense, are to be supposed to be. Thus, in an everyday sense we might say that if it is a norm in some context C to φ, then this means that to φ in C would be correct or appropriate, and if it is a norm in C to not φ then it would be incorrect or inappropriate in C to φ. And *then* it might be asked whether the norm is social and dynamic, how people learn about and understand the norm, what acting accordingly involves, and whether it applies necessarily or contingently to the specific domain.

The meaning of 'correctness' and 'appropriateness' here can be understood in terms of what position one puts others and oneself in by the deed in question. Acting incorrectly permits others to demand correction or to demand a reason why one thinks that one is not committed to correct

oneself. Perhaps it might also lead to sanctions, if that is a norm. Acting correctly might not entitle to praise or admiration, but it entitles to not having to correct oneself.

So the strategy adopted in this dissertation is not that of stating a general definition of 'norm' in order then to investigate various instances of norms and how they are learned, understood, and acted in accord to. Instead each paper gives its own, specific, and domain-relative tack on the issue of what a norm is (expect in papers 2 and 3, which treat social understanding). In this respect, perhaps the meaning of 'norm' should be thought of in terms of Wittgenstein's notion of *family resemblance*. He writes (1958, p. 17) concerning the meaning of 'game':

> craving for generality is the resultant of a number of tendencies connected with particular philosophical confusions. There is—(a) The tendency to look for something in common to all the entities that we commonly subsume under a general term.—We are inclined to think that there must be something in common to all games, say, and that this common property is the justification for applying the term "game" to the various games; whereas games form a family the members of which have family likenesses. Some of them have the same nose, others the same eyebrows and others again the same way of walking; and these likenesses overlap. The idea of a general concept being a common property of its particular instances connects up with other primitive, too simple, ideas of the structure of language.[3]

This reminds also of von Wright's warning against a General Theory of Norms. Still, and to the contrary, I believe that the background provided here, together with the clarifications in the papers, will counteract the impression that we do not know what we mean by 'norm' across contexts. This is so because although what we mean by it in one area of research may be somewhat different from what we mean by it in another, still there seems to be a core everyday sense which we may cling to so as not to have to redefine the term completely and anew in each case.

Let me end this introduction with positioning my understanding of what an analysis of 'norm' must ultimately be by means of referring to another contemporary philosopher whose work has greatly influenced parts of this dissertation.

1.3 Norms and Analysis

I want to give one further motivation why I do not attempt what von Wright called a General Theory of Norms, or an analysis of the concept that generalizes across the whole field. I do this by means of a specification of what an analysis of norms, as I see it, is: it is itself normative activity.

An analysis tells us *how to understand* the analysandum. As such, one of its points is to tell us: it is correct to say certain things, according to the analysis, about what is analysed; other things are incorrect. A conflict between analyses can thus be understood as a conflict on how we should understand some phenomenon such that we should and should not say certain things about it. This is what I mean by an analysis of norms itself being a normative activity.

My reason for not attempting the General Theory is that I want instead to understand how people learn norms and act according to norms in specific circumstances (which might, but does not in this dissertation, involve also the question how people learn norms for analysis and analyse accordingly). So I distinguish the question what an appropriate analysis of the term 'norm' is (which is not the question with which I will be concerned, more than summarily) from the question what is involved in learning, understanding, and acting as is normative in specific situations (which is the question I will be concerned with).

It might then be asked if what I am doing here is not also normative. I believe it is. I happily concede that this dissertation is a normative contribution to the debate how to correctly understand what is involved in people learning, understanding, and acting according to norms. That is, I say something about what is correct to say about people's understanding of and action according to norms.

My approach to an analysis of norms can be illustrated by an example from Robert B. Brandom. He discusses the circumstances and consequences of introducing a new term in a language, or of introducing a new sense of an old term, and argues that these are essentially normative. Brandom (2000, p. 69, quoting Michael Dummett 1973, p. 454) writes:

> The conditions for applying the term ['Boche'] to someone is that he is of German nationality; the consequences of its application are that he is

> barbarous and more prone to cruelty than other Europeans. We should envisage the connections in both directions as sufficiently tight as to be involved in the very meaning of the word: neither could be severed without altering its meaning. Someone who rejects the word does so because he does not want to permit a transition from the grounds for applying the term to the consequences of doing so. The addition of the term 'Boche' to a language which did not previously contain it would produce a non-conservative extension, i.e. one in which certain other statements which did not contain the term were inferable from other statements not containing it which were not previously inferable.

One way to interpret Brandom here is as saying that we must always be on our guard against the normative circumstances and consequences of definition. What is proposed as *descriptions* often has *normative* circumstances and consequences of application. Brandom notes that this is true of many words such as 'faggot', 'Communist', 'nigger', 'whore', 'lady', etc.[4] How we define these, what we accept as entitling us to use them in some circumstances and in what circumstances we are committed to apply them, reflects our readiness to endorse certain norms.

Consider an excerpt from Oscar Wilde's trial. The judge reads aloud from a passage Wilde has written and says, "I put it to you, Mr. Wilde, that this is *blasphemy*. Is it or is it not?" Brandom (ibid) writes:

> Wilde made exactly the reply he ought to make—indeed, the only one he could make—given the considerations being presented here and the circumstances and consequences of application of the concept in question. He said, "Sir, 'blasphemy' is not one of my words".

If in Wilde's community the consequences of conceding to blasphemy were taken to entitle others to, e.g., accuse one for being a criminal, which in turn were understood to be sufficient conditions for a judge to be committed to issue a penalty, then Wilde's best option would, indeed, be to say that blasphemy was not one of his words; i.e., he did not accept to take what he had written to entitle the inference to him being a criminal, and so on. We can say that what Wilde implicitly did in giving the response he gave was to contest that what he had written provided a circumstance that entitled and committed to certain consequences, such as accusing him for being a criminal. Now, compare this story to how we are to understand 'norm'.

Asked whether or not something is a norm I might say: 'norm' is a contestable concept. In the context of any particular situation it is contestable whether some activity is normative (I contest that it is not contestable!). To say that something is a norm is to take a normative stance, because an implicit consequence of doing so is to license entitlements and commitments to judge and hold others responsible to it in certain circumstances. It might be true that something is a norm in my community and hence it might be true that I cannot do otherwise and be said, relative to this community, to act correctly. But one may always contest that something is a norm as well as contest the meaning that is attached to 'norm'. In any case, 'norm' is not a word whose meaning I am primarily interested to contest here (what I would contest is that I am not entitled to contest it). Instead I am concerned with the question what we mean when we say that someone *does* something; namely, when she understands something to be a norm and when she acts on her understanding of something to be a norm.

Norms permeate human activity, from standing distance in conversation to philosophical analysis. As for analyses of 'norm', such analyses are themselves normative; they are efforts, in an ongoing normative activity of telling us how we should understand what a norm is, to say that we should in all cases understand the term in one particular manner rather than another. Both the suggestion that we should under-stand the term in one particular manner and the claim that that meaning should be taken to be the same under all circumstances are normative claims. I put it to you, the reader, that this dissertation is not an analysis of 'norm' but an inquiry into how you, others, and I understand norms and what we are doing when we act accordingly. That is how it is correct to understand how this investigation is done.

Notes:

[1] Another area in which norms are often claimed to be essential is reasoning and rationality (Zangwill 2005; 2010; Wedgwood 1999; 2002). For example, it might be argued that the reason why we understand a propositional attitude as being in the psychological mode, e.g., belief, that it is, is because we recognize it as normatively requiring the rejection or acceptance of certain other propositional attitudes. That I *believe*, e.g., that it is raining and that if it is raining

then the streets will be wet, means that I am normatively committed to *believe*, and not merely, e.g., *guess*, that the streets will be wet, quite regardless of whether the antecedent beliefs are true.

We may also speak of norms for practical reasoning and instrumental rationality. For example, it might be argued that to infer an intention to get on the train, or to just act, is normatively required if one believes that the train is about to leave and desires to get on the train. In this sense one has a reason to get on the train and the reason, it may be argued, is a norm that demands action.

In short, there is an exciting flurry of suggestions that norms apply, necessarily or contingently, in this or that domain and level of conduct, many of which I have not had the opportunity to investigate in this dissertation.

[2] Kripke's interpretation of Wittgenstein's reasoning is one of the major culprits for the vigorous and exciting three decades-debate on whether meaning is indeed normative, from which parts of this dissertation takes off. It should be mentioned that there are several other interpretations of Wittgenstein that I have not had the opportunity to adequately relate to. See, e.g., Wright (1980), McDowell (1984), and McGinn (1984).

[3] For an interesting discussion on the coherence in Wittgenstein's own use of the term 'family resemblance', see McGinn (2011, chapter 2: "Definition and Family Resemblance").

[4] Brandom's position (see also his 1994 and 2008) is, roughly, that conceptual content in general is essentially normative in the sense that it is to be defined in terms of committing and entitling to inferential relations that people in social discursive practices implicitly take to apply. For instance, 'red' is to be understood in terms of implicitly committing someone saying "*x* is wholly red" to endorse the proposition "*x* is coloured" and as not entitling to "*x* is green". That commitments and entitlements are implicit here means that the person may not herself be in position to state that she is so committed and entitled. The same can be said about the terms 'norm' and 'normative'; i.e., in using them one implicitly commits and becomes entitled to certain inferences that one might not also explicitly state that one is committed or entitled to.

2. Theory, Definition, and Method

I approach the study how people learn, understand, and act in accord to norms against a background of two competing models. These are the *cognitive* and *non-cognitive* models. In this chapter I specify the two (sects. 2.1-2.3) and define central terms not clearly defined in the papers that follow. I then discuss the method of this dissertation in relation to how it approaches the question how people learn, understand, and act in accord to norms (sect. 2.4), given how this is understood on the models discussed.

2.1 Background On Models: Cognitive and Non-Cognitive

There are two main approaches to the study of how people learn and act in accord to norms: *cognitive* and *non-cognitive* models.[1] Discussing these will provide the theoretical background to this dissertation.

On the first, cognitive model, understanding is theory-like in structure, as follows. One's understanding of others or how to act has a propositional and inferential form such that, to understand another or an activity as normative, one is required to derive a belief that something is correct or that someone is in a specific state of mind from an observation of her behaviour, and infer, from this belief, what to do, or what mental state to attribute to the other, according to certain rules of inference. For example, that Sally understands that three feet is the correct standing distance in conversation, and her acting accordingly, involves: Sally believes a proposition stating that three feet is the correct standing distance in conversation; she believes

that this proposition is true and applies to her current situation; her beliefs are true; she infers from these premises that she should now keep approximately three feet distance to her interlocutor. She concludes by intending to position herself accordingly. If this is the case then Sally can be attributed an understanding of the norm for standing distance in conversation. If she acts in according to this understanding then she is said to act in accord to the norm.

The cognitive model, then, requires that people entertain certain propositionally contentful states *that* something is a norm. From these it is necessary to infer an intention to act in accord to the norm. In other words, understanding norms involves having a theory about correctness, constituted by a set of true beliefs and inference rules for deriving true propositions about correctness in the situation at hand. Acting according to norms is to act on an intention outputted by the process of running the theory.

On the second, non-cognitive model, understanding is not theory-like in the above sense. Instead, it is a kind of skilfulness, a technique to respond to norms in ongoing activity. This understanding might not be propositionally articulated and attained by means of rational inference. It is rather a form of knowledge *how* to act, opposed to knowledge *that* something is correct, as emphasized on the cognitive model.

This dissertation is in large a contribution to a non-cognitive model of how people learn and understand norms, and act accordingly. But the non-cognitive model comes in many flavours, each reflected in a different way in the papers that follow. So let me say something about these different flavours here. In so doing I provide definitions of central terms, definitions that are lacking in the papers. In a later section (2.3) I will have more to say about the cognitive model.

2.2 The Non-Cognitive Model and Definitions

In this section I specify the variants of the non-cognitive model and define central terms.

It should be noted here already that the terms that I use (embodiment, situatedness, enaction, dynamics, and ecology) have a fairly short history,

especially in the context of analytic philosophy, which is the context of this dissertation. As of today there is still debate among proponents of a non-cognitive model how precisely to understand the keywords. Therefore, what I say in what follows in this section must be read cautiously; what I say may not yet be generally agreed on. Nonetheless, I intend to give enough space for definitions of several fairly ingrained positions, so as not to be accused for missing out on any that might be argued to have survived long enough to be a major player in the field. It is important to keep in mind, though, that the following fields of research are burgeoning and that progression and change is happening even as this sentence is written.

2.2.a Embodiment

On a first version of the non-cognitive model, understanding is said to be *embodied*. This can in turn be understood in different ways.

It might first of all be asked in opposition to what kind of 'disembodied' approach understanding is embodied. After all, it is difficult to find anyone claiming that we are not embodied in this sense at least: we have bodies.

What we might call the disembodied approach does not claim that we do not have bodies but that the nature of our embodiment has no implications for the nature of our understanding. As Clark (1999, p. 347) formulates it, the disembodied view can be characterized by it "positing an inner realm richly populated with internal tokens that [*stand*] *for* external objects and states of affairs", where this "inner realm" is insulated and causally independent from bodily states and embodied activity. If the claim of embodiment approaches were simply that we have bodies, then it would not stand in opposition to this understanding of embodiment. But the interesting claim is that a specification of the nature of our embodiment (in a weak, moderate, or strong sense; see below) has consequences for a specification of the nature of our understanding. Indeed, on one version of this claim, embodied action *constitutes* understanding. It is against the view that embodied activity neither contributes to or itself is part of understanding—what we might call a disembodied approach—that embodiment theorists revolt.

Let me now distinguish three senses of embodiment.[2] On a first construal, understanding is embodied in the *weak* sense that one's understanding is influenced by what possibilities for action and perception comes with having the body one has. Morphological and structural properties of one's body, such as the number and arrangement of cones and rods in the retina and the elasticity of tendons at the joints, constrain and enable certain ways of perceiving and interacting with one's world and with others, and this constrains and enables certain ways of understanding others and the environment. However, understanding is itself decoupled from one's embodiment, in the sense that it is attained by means of interpreting what one perceives such that one can acquire a representation of the world, which then informs one's activities. Hence on this weak sense of embodiment one's body functions only as a dead input-output device for reaching an understanding and to inform action accordingly. We may specify this weak sense, what Kiverstein (2012, pp. 740-41) calls "body conservatism", as follows:

> *Weak Embodiment:* Embodied action supplies perceptual input for disembodied symbol manipulation and for execution of motor instructions derived from such central symbol processing.

From the point of view of weak, or what Clark (2008a, p. 42) calls "mere", embodiment "the body is nothing but a highly controllable means to implement practical solutions arrived at by pure [disembodied] reason" (ibid). Insofar understanding is concerned, embodiment, in the weak sense, is only important inasmuch as it produces input to a central, disembodied, reason, and consumes instructions from this reason in a trade-off for behaviour production. The body is, considered in itself, a physical, inert system, a corpse, awaiting the management and control of a reason whose embodiment is of no consequence for it.

In a second, *moderate*, sense of embodied the body is said to *functionally* participate in the structuring of one's understanding. Being embodied in this sense is not only to have a physiology allowing one to perceive and act. Embodied activity structures one's perception and action and thus it also shapes one's understanding—how what is perceived is represented and how representations in turn inform embodied action that

in turn cause updates for further perception. Embodied action, a kind of active exploration of possibilities for interaction in an environment, including others, is itself part of reaching an understanding how to act. And embodied action might transform the environment, including others' actions, which provides further possibilities for action. In this loop of acting and being acted on, understanding is constantly formed and transformed in embodied activity. Thus one's embodiment plays an important role in any explanation of the contents and form of one's understanding. Nevertheless, the body is on this view a *functional* platform; it matters "only insofar it plays the right functional role" (Wilson and Clark 2009). We may specify this moderate sense of embodiment, what Kiverstein (2012, p. 740) calls "body functionalism", as follows:

> *Moderate Embodiment:* Embodied activity plays a functional role in how the subject understands how to act and in which contingencies to respond accordingly.

On the moderate understanding of embodiment two subjects with biologically and physiologically different bodies could, in principle, have the same or at least a very similar understanding of how to act, given that their bodies are functional equivalents. This view "lends no support to the idea that minds like ours require bodies like ours" (Clark 2008a, p. 203), because bodies physiologically and biologically *un*like ours *could* implement the same functions as ours.

On a final, third, and *strong* sense of embodied, one's embodiment is said to be *constitutive* of understanding. Here it is not as a functionally defined living platform that one's embodiment matters for understanding, but rather as phenomenologically *lived* (Thompson 2007). This means that embodied action *is* a kind of understanding that is also experienced as an ongoing activity. Whereas on the moderate construal embodiment is in principle platform-neutral in the sense that one's understanding could be the same despite one's embodiment having different physical and biological properties insofar as it implements the same functional properties, differences in physical and biological embodiment does, on this strong construal, make a difference for one's understanding and experience of one's activities, environment, and others. This is what it means to say that the

body is *lived*. Having the experiences that comes with having a certain body implies also making sense of one's world in a specific manner, a manner not transferrable to a different albeit functionally equivalent body. This strong sense of embodiment has been called "radical" "body-enactivism" (Kiverstein 2012, p. 741) and connects to the idea (more on which below) that embodied action is *creative*; that it is constitutive of what the subject comes to understand and the process in which this is achieved. We may specify it as follows:

> *Strong Embodiment:* Embodied action is the source of understanding and is constitutive of how the subject understands how to act.

Indeed, on this view the very thermodynamic and biological processes that comes with having a particular kind of body is intrinsic to having the experiences and understanding one has. For this reason functional equivalence is not sufficient for two subjects to have similar experiences and understanding, because this requires having the same biological constitution.

This dissertation contributes to an understanding of norms as realized, identified, and understood in embodied action. Unless otherwise specified I will mostly be concerned with investigating the hypothesis that understanding is embodied in the strong sense. That is to say, I will be interested in the proposal that an understanding of others and norms that apply to some practice would be different if those who participate in the practice were differently embodied not merely functionally but—in the stronger sense—even thermodynamically and biologically, where this is understood to be constitutive of subjects' experience and understanding. This claim holds that people would understand others, their environments, and how to act differently if their bodies instantiated different biological and thermodynamic properties. In general, animals with the same biological and thermodynamic nature should with more ease be able to understand each other and be more likely to share norms than if their bodily constitution were different, because their experiences and understanding would be similar in the first case and different in the latter.[3]

One reason for my investigating the strong sense of embodiment in particular is that it is the sense of the term that, if found necessary to understand people's understanding and acting according to norms, most

forcefully opposes a cognitive model. To claim that embodied activity constitutes understanding and not only contributes to it is more of a challenge to the claim that understanding is not embodied because it effectively replaces the notion of some disembodied understanding to be contributed *to* with the notion that all understanding is embodied—i.e., there is no understanding somewhere in the body, separate from embodied activity, that embodied activity causes or is caused by.

Another reason for investigating the strong sense is that a big question in contemporary embodiment research is whether the strong sense is an improvement on the other two and if it is compatible with them. Now, although I will not get involved in that debate specifically, an answer to the question whether understanding others and norms is strongly embodied will expand on the number of issues that should be taken into account. For example, if it turns out that our understanding of norms must be understood as strongly embodied, or if it cannot be so understood, then this makes for one particular contribution to the debate whether the nature of human understanding *in general* is strongly embodied. For suppose that to understand people's understanding of norms *requires* an understanding of their biological and thermodynamic embodiment. In this case advancement would have been made on whether the nature of our understanding is strongly embodied in general, by finding a particular case in which it is. So although I am not favouring or arguing for one specific definition of embodiment, I do favour a strategy of investigating the strong sense because more hinges on the outcome of that investigation than on an investigation of the two weaker notions. Hence when I speak of embodiment I intend the strong sense, unless otherwise stated.

2.2.b Situatedness

This dissertation also investigates a 'situated' approach to how people learn norms and act accordingly.[4] The situated approach (e.g. Gallagher 2004) insists that rather than understanding being propositional and theory-like in structure, as on the cognitive model, it consists in a practical know-how to directly, pre-reflectively, read-off from social environments and others' actions what is correct, and how to act accordingly. In this sense, the situated version of a non-cognitive model expands outward what a theory of

understanding in general should take into account. From focus on embodiment to focus on situatedness lies a step of expansion with respect to what is relevant for the nature of a subject's understanding. This step should not be taken *away* from embodiment in the sense that embodiment is no longer to be considered important, but as a follow-up along a continuous non-cognitive line of thought.

On the situated approach people are supposed to interact with each other in pragmatic contexts; i.e., in environments defined in terms of social roles and statuses people have in their community and what proprieties for action comes with such roles and statuses. Aron Gurwitsch (1931/1978, p. 108) writes:

> The originary encounter with fellow human beings does not signify a coming together and being together of isolated individuals who, in their mutual encounter, have severed their collective relations to the surrounding world and, so to speak, find themselves together but detached [...] as *mere individuals* such that this sort of encounter would be a *mere being together*. Instead, we continuously encounter fellow human beings in a determined horizon.

This "determined horizon" is the social realm of roles and statuses that people are entangled in by virtue of the very fact that they participate in the social practices of a community. The other is encountered *first* as someone with socially significant properties, which may *then* be cognitively abstracted from by stripping away from the encounter what primarily and prereflectively matters to us (ibid, p. 35). Thus, for example, being a police officer comes with certain proprieties for action. The same is true for being a teacher, a Professor, a priest, a spouse, and so on. A social encounter is always, on this view, an encounter with someone in his or her role. And thus one encounters the other as someone subjected to certain proprieties for action. These might also have consequences for oneself, in one's own role. For instance, the permissions and prohibitions that define someone's role as a police officer, spouse, Professor, etc., might have implications for what others, e.g., a lover, arrestee, student, etc., are permitted to or prohibited from doing in interaction with him or her. And so, encountering the other in his or her social role or status is also an encounter situated in a

network of norms applying to oneself and others and to interactions among people in general.

As I investigate this approach, appropriate procedures for interaction are found to often be 'indicated' (Searle 1995) in social and physical environments such that people suitably situated are in position to directly read-off from them how to act appropriately in a particular context. Among indicators we find badges, uniforms, wedding rings, the physical layout of a lecturing hall, title deeds, sirens, and so on. What defines the class of indicators is that their function is to draw people's attention to something about the role or status of people or objects in the context of particular situations in which they interact. For example, to see a person in a certain uniform is a reliable indicator that this person is a police officer. And insofar being a police officer in the relevant community comes with certain permissions and obligations, both of police officers and those who interact with them, seeing someone in police uniform is a reliable indicator that there are certain appropriate manners for this person to act and for others to interact with her. This means that, on the situated approach, one can come to understand what norms apply in a particular context by being attentive to indicators in it rather than by interpreting others' behaviours and inferring what mental states they must be in, in order thereby to conclude how to act.

To provide a definition of the situated approach as a unified theory might be somewhat misleading, because there are many different formulations of what is meant by people having situational understanding. However, I would specify the approach that I will investigate as follows:

> *Situated Understanding:* That understanding is situated means that it is enabled and partly constituted by contextual factors in physical and social environment in which it occurs.

As is true with respect to embodiment, also situatedness might be given weaker and stronger interpretations. On the "extended mind" theory (Clark and Chalmers 1998; Clark 2008a) the claim can be put as follows: environmental structure—artefacts such as notebooks or maps—are *constitutive* parts of processes of understanding in the sense that manipulation of that structure is itself a process of embodied action that can be characterised as constitutive of understanding. For example, if my

looking up an address in my notebook is part of the process that culminates in my understanding where to go and how to get there, then the notebook and my manipulations (looking things up in it) can be conceived of as part of my understanding to the same extent that neuronal firings would be part of my understanding if I were bereaved the external aid (Clark and Chalmers 1998). On a weaker version the claim might be that social and physical environments *causally contribute* perceptual input, which is processed only internally (in the brain) in the form of symbol manipulations (Adams and Aizawa 2010a). In this case, to say that understanding is situated would amount to saying that environments, like embodiment on the weak interpretation, are only something responded to and acted in but not themselves constitutive of, or partaking in the shaping of, understanding. With respect to this, the thesis I will investigate is that understanding is often *strongly* situated. This means that in many cases an individual's understanding of another or of how to act appropriately constitutively depends on her social and physical environment. If the social and physical environment were different the individual would understand others and the situation's normative requirements differently, or not at all, *not* because of something internal to the individual but because of her whole, situated, comportment.

This specification brings into relief the contrast between this version of a non-cognitive model and a cognitive model. On the situated approach, an individual can have an understanding of how to act appropriately that does not require her to know anything about the particularities of the other's mind—intentions, beliefs, desires—because in being situated in a social environment she can come to directly understand what people, the 'we' that constitute the community, take to be appropriate. So her situatedness in a wider social context might 'short-circuit' the theoretical inferences about other minds that on a cognitive model would be required for her to understand others and how to act. We may say that this emphasizes the *socialized* nature of understanding—it being constituted by socially contextualized encounters—in contrast to it being a *psychologizing* in the sense that Gallagher (2008, pp. 535-36) opposes, namely:

> that the normal and pervasive way in which we understand other persons depends on a practice [...] in which we employ common sense or

> folk psychological theory about how mental states (beliefs, desires, intentions) inform the behaviors of others [that] allow us to explain or predict the other person's behaviour or what they must be thinking.

On the situated approach, psychologizing in this sense is precisely a secondary act of cognitive abstraction from, in Gurwitsch's terms, the "originary encounter".

It should be noted that the situated approach that I investigate here does not, if vindicated, warrant the strong conclusion that understanding is *never* theory-like. It is rather an antidote to the claim—variously known as the "intellectualist legend" (Ryle 1949/2009, p. 18) and "the myth of the mental" (Dreyfus 2007)—that understanding what to do is *always* theory-like. If in some cases understanding is strongly situated, then it is not true that it under all circumstances involves theoretical reasoning along a observation-interpretation-representation-inference-action sequence. And if the conclusion that it does not is warranted, then this shows that a cognitive model of the nature of our understanding must be somewhat restricted in its scope; at least with respect, insofar this dissertation is concerned, to the nature of understanding others and norms.

The reason why I focus my investigations on the strong sense of situateedness is similar to the reason why I focus on the strong sense of embodiment. The reason is that it is the strong sense that would, if supported, most clearly bring out the shortcomings of classical understandings of the nature of our understanding as something in principle abstracted from a wider context of embodied and situated activity.

2.2.c Ecology

The ecological approach to normativity that I will be concerned with, as part of a non-cognitive model, is an elaboration of James J. Gibson's (1979) original formulation of an ecological approach to visual perception.

The original approach is an attempt to understand visual perception as a process of functionally embodied agent-environment interaction. The idea is that an agent's perception of its environment is a function of its embodiment in conjunction with the structural properties of the environ-

ment. So what you see is function of what you can do, and vice versa (recall the moderate sense of embodiment).

Gibson labels the relations that obtain between agents and their environment as a function of their embodiment and the environment's properties, which he argued are what agents are directly perceptually presented with, 'affordances'.[5] He writes (ibid, p. 127):

> The *affordances* of the environment are what it *offers* the animal, what it *provides* or *furnishes*, either for good or ill. The verb *to afford* is found in the dictionary, but the noun *affordance* is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment.

Gibson thought of a kind of animal's environment as a 'niche'. The niche of the environment of a kind of animal he describes as, not *where*, objectively, it lives or how it subjectively *experiences* its environment, but functionally as *how* it lives (ibid, p. 128). He thought that this functional description of an ecological niche of affordances could dispel the traditional dichotomy of perception being either of an external objective reality as described by physics, on the one hand, and of it being something internal and subjective to agents as this is described in psychology, on the other. In place of this dichotomy Gibson wanted to say that perception is primarily of affordances, and these are neither objective nor subjective but relations cutting across that divide because they consist in the ecology of kinds of animals with their particular bodies and interests and the features of the environment that relative to the kind of animal offers certain possibilities for action.

In my investigations on how people learn and understand norms the concepts of affordance and niche are put to the task to illuminate whether people may not also be directly responsive to norms considered to be affordances of their shared niche. I speak of niches of social practices. These niches are to be thought of as consisting in patterns how people interact socially with each other, which over time establish what we may call customs of communities (Rietveld and Kiverstein 2014, pp. 328-29) or of smaller subgroups. A niche of social practices, then, not only specifies how people live relative to their embodiment and the environment, but also how people live *together* and interacts socially in a shared environment.

Already Gibson noted that what animals afford each other is not only behaviour but also social interaction (1979, p. 42). I have been interested in the notion that what a niche affords its inhabitants is not only a *functional* way to interact with each other but also a *normative* way to interact *correctly*. So on the ecological approach to norms, norms are not specified relative only to attitudes internal to the psychology of any agent considered in isolation from a niche of social practices. Nor are norms external and objective, specified as something part only of a shared physical and social environment. Instead they are ecological relations (affordances) that people establish together in time in social interaction, such that they become part of a niche, the affordances of which people directly respond to.

In line with the ecological approach to visual perception, an implication of an ecological approach to norms and people's understanding of norms would be that norms are what people *first* respond to and is not something they must interpret as exemplified in, and attribute to, a social practice in order then to infer what to do. It is the other way round. People may be understood to directly understand how to act correctly by way of participation in the social practices of a niche and secondarily to learn to infer and state in propositional form that thus-and-so is a norm. Thus affordances considered to have normative significance are not only for action or for social interaction but also for correct social interaction.

So what is meant by an ecological approach to norms and understanding of norms can be specified as follows:

> *Ecological understanding:* That understanding is ecological means that it consists in a capacity to discriminate correct from incorrect social interaction afforded relative to a niche of social practices.

Now, if the situated approach is an extension of focus in research on social and normative understanding from an embodied approach, which is an extension from focus on disembodied and abstract reasoning, then the ecological approach is a unification of the embodied and situated nature of understanding. What is correct in a niche of social practices implies both a particular embodiment of its inhabitants as well as a particular manner in which they are situated. What is normative, on this view, "implies the complementarity of the [animal's embodiment] and the environment [in

which it is situated]", to borrow Gibson's words (ibid, p. 127). And hence also to understand how to act correctly in the way that we do would require being embodied like we are and situated in a niche of social practices like ours. Here 'we' and 'our' refer to social contexts of wider or narrower scope, ranging from humans in general to particular socio-cultures, with a corresponding alteration in degree of probability that people would understand each other, as a function of differences or similarities in the embodiment and situatedness that constitute their niches.

I have described the cognitive model as committed to the claim that understanding is an achievement of theoretical reasoning according to a sequential input-representation-inference-output structure. If people's understanding of norms can be understood as ecological, i.e., as direct, practical, and constitutively dependent on embodied activities relative to an environment, including others, then the cognitive model is revealed to posit, as necessary for social and normative understanding, processes that are *not* necessary.

The motivation for investigating the possibility to understand norms as affordances is thus in line with the general motif of this dissertation; namely, to inquire into the prospects for providing a non-cognitive model of the nature of our understanding each other and norms. The ecological approach fits the picture because its vindication would provide momentum for a non-cognitive model of the nature of human understanding of norms and normative activity.

2.2.d Enaction and Dynamics

Enactivism is a theory in the philosophy of the cognitive sciences and its opus classicus is Francisco J. Varela et al.'s *The Embodied Mind* (1991)—although it traces back at least to Humberto R. Maturana and Varela's less attended to *Autopoiesis and Cognition: The Realization of the Living* (1972/1980).

Central to the enactivist approach[6] is that cognition is embodied. On the enactivist view that I will be investigating a strong sense of embodiment is favoured. The very biochemical and thermodynamic processes of one's lived body is claimed to be constitutive of cognition (Thompson and Stapleton 2009). To this, enactivists add that cognition is 'enactive'. This

means that cognitive beings have the mind and understanding they do as a consequence of dynamic processes of interacting with each other and their environment. This process is characterized as dynamic because it constantly produces non-linear change in patterns of thinking and interacting, which, in turn, produces new and change already established patterns (recall sect. 1.1.a). On-going embodied activity, situated in a physical and social environment, is an *active* or *creative* process in the sense that it constantly re-constitutes and transforms understanding. Understanding others and how to act is a skill or technique to actively negotiate dynamic change, in contrast to a capacity to interpret behaviour and follow rules conceived as linear specifications of the correct way to go on given a set of initial conditions. As Varela et al. (1991, p. 148) put it:

> commonsense knowledge is difficult, perhaps impossible, to package into explicit, propositional knowledge—"knowledge *that*" in the philosopher's jargon—since it is largely a matter of readiness to hand or "knowledge *how*" based on the accumulation of experience in a vast number of cases. [...] Indeed, if we wish to recover common sense, then we must invert the representationist attitude by treating context-dependent know-how not as a residual artifact that can be progressively eliminated by the discovery of more sophisticated rules but as, in fact, the very essence of *creative* cognition.

And, they continue (ibid, p. 149):

> The central insight of this [...] orientation is the view that knowledge is the result of an ongoing interpretation that emerges from our capacities of understanding. These capacities are rooted in the structure of our biological embodiment but are lived and experienced within a domain of consensual action and cultural history.

I will investigate an enactive approach to social understanding in particular. The germ to this approach is quite new (e.g., De Jaegher and Di Paolo 2007). I have therefore had to engage in a piece of pioneering. On the enactive approach that I investigate 'enactive' should be understood (cf. De Jaegher et al. 2010) as follows:

> *Enactive understanding* That understanding is enactive means that it consists in a dynamic process of embodied participation in social interaction in which two or more subjects mutually negotiate how to act in the context of their particular interaction.

Enactive understanding is thus highly context-sensitive, because social interaction as a dynamic process constantly changes and establishes norms in the particular situation engaged. Social interaction is here also understood in a strong sense, meaning that no one individual is unilaterally managing the process but that all participants together, mutually and reciprocally, shape it (Froese and Gallagher 2012). If one individual is directed or controlled by the other(s) then this means, on the enactive approach, that that individual is not a *participant* to the interaction. This is what is meant by "participatory sense-making" (De Jaegher and Di Paolo 2007; Fuchs and De Jaegher 2009). Each participant's activities modifies the progress of social interaction, which in turn modifies each participant's understanding of the situation (De Jaegher and Di Paolo 2013, p. 1), where this happens despite no one individual knowing or intending that the process progresses as it does. The progression of social interaction and the social understanding achieved constitutively depend on the autonomous contribution of all participants together, not on any one considered in isolation from it.

And so, to say that people's understanding of each other, as well as of the meaning of language and how to act more generally, that it is *enactive* means that it is constantly in a process of change in social participation. Understanding is a dynamic process of knowing how to act that is both caused by and causally involved in progression of embodied social interaction that produce and change norms and meaning of behaviour in particular contexts.

Engagement in social practices—actual embodied participation in the presence of the concrete context of such practices together with others—is, on the hypothesis that understanding others and norms is enactive, required for acquiring a know-how to act appropriately. Hence, enactive understanding is, quite literally, socially engaged and participatory, in the sense that it consists in the ongoing embodied interaction with others where no one individual would reach the understanding on her own. De

Jaegher and Di Paolo (2007, p. 497) describe the participatory nature of social understanding as follows. It is,

> the coordination of intentional activity in interaction, whereby individual [understanding] processes are affected and new domains of social [understanding] can be generated that were not available to each individual on her own.

As I understand it, enactivism is the natural next step from embodiment, situatedness, and ecology, in that it can be seen to build on these but also as extending the story with a chapter on our situated and ecological understanding being enacted and reciprocally enacting ever new situations and ecological niches for further possibilities of participation and understanding. If plausible, it can help us understand how norms are established in processes of embodied participation in social practices and how these constantly establish new and change existing norms.

Enactivism thus braids together situational and embodied approaches, and may constitute one answer to what norms are, how they evolve, and how we understand them. What it suggests is that meaning and norms are highly context-sensitive and that therefore people's understanding and normative acitivity must be specified relative to the particularities of the context, or indeed, niches, in which these abilities are exercised and developed.

The interest in an enactive approach is again a testimony to the general hypothesis that this dissertation examines: whether a non-cognitive model can plausibly tell us how people understand each other, meaning, and norms, and what it is that they are doing when they act according to such an understanding.

2.2.e Summary

In summary, it should be noted that the four versions of the non-cognitive model follow a common line of thought. For instance, embodiment plays a central role, to various degrees and with different interpretations, for situated, ecological, and enactive understanding. Likewise, the niche of the ecological approach has certain similarities to that of a context of interaction for creative cognition emphasized on the enactive approach. And

understanding in the context of a niche or context of enactive cognition is *eo ipso* also situated and embodied.

Therefore, although there may be disputes between as well as within these four approaches on how to understand central terms, they can be seen to agree on and mutually support the general claim: our understanding of each other and norms, and our capacity to act accordingly, is to be understood as necessarily embodied, socially situated, ecological, and dynamic. This dissertation is meant to contribute to an understanding of this claim.

I have not yet said much about the competitor cognitive model. I do so next. After that I present the method employed in the papers that follow.

2.3 The Cognitive Model

One popular idea is that norms require for their existence propositionally articulated cognitive states like belief, recognition, acceptance, or expectation *that* something is collectively taken to be correct or appropriate. For example, that it is a norm not to bike in the park might be thought to entail that people collectively accept that the proposition that biking in the park is incorrect has directive force (do not bike in the park!), that people believe that others believe this, that people expect each other to not bike in the park and that, if one bikes in the park, others will think unfavourably of one, and perhaps sanction one's behaviour.

The understanding of norms here fits a cognitive model because it requires beliefs, expectations, and other cognitive states, in a population whose members interpret and infer each other's cognitive states in order to infer true propositions about what is the correct thing to do. It is a process with an observation-belief-inference-behaviour structure.

For an illustration of this approach we may consider David Lewis's version, from his book *Convention* (1969). He analyses norms as "regularities to which we believe that we ought to conform" (p. 97).

According to Lewis (ibid, p. 42), a convention is to be defined as follows:

> A regularity *R* in the behavior of members of a population *P* when they are agents in a recurrent situation *S* is a *convention* if and only if, in any instance of *S* among members of *P*,
>
> (1) everyone conforms to *R*;
> (2) everyone expects everyone else to conform to *R*;
> (3) everyone prefers to conform to *R* on condition that others do, since *S* is a coordination problem and uniform conformity to *R* is a proper coordination equilibrium in *S*.

Now if *R* is a convention in a population *P* in recurring situation *S*, then (ibid, p. 97): (1) most members of *P* will conform to *R* in *S*; (2) an individual prefers to conform to *R* in *S*, given that most others in *P* do so; (3) most members of *P* have reason to expect others to conform to *R* in *S*; (4) most members of *P* prefer to conform to *R* in *S*, given that others conform; and (5) each member of *P* has reason to believe that (1)-(4) hold. This means that members of *P* have 'common knowledge', i.e. that everyone has reason to believe that (1)-(4) hold, which indicates to everyone that everyone has reason to believe that (1)-(4) hold, which indicates to everyone that everyone has reason to believe... and so on.

If I am in this kind of scenario, then I have reason to believe that my conforming to the regularity would satisfy my own preferences as well as the preferences of others who have reasons to expect me to conform to the regularity, which is something I have reasons to believe and they have reasons to believe that I have reasons to believe that they have reasons to... and so on. Moreover, others have reasons to believe that my conforming to the regularity in the situation with which it is associated would satisfy my own preferences. Hence others also have reasons to believe that I have reasons to believe both that my conforming will satisfy their preferences and that they have reasons to expect me to conform since my doing so would satisfy my own preferences. If all of this obtains, then (ibid, p. 99):

> if they see me fail to conform, not only have I gone against their expectations; they will probably be in a position to infer that I have knowingly acted contrary to my own preferences, and contrary to their preferences and their reasonable expectations. They will be surprised, and they will tend to explain my conduct discreditably.

The idea here is that for a behavioural regularity to be a norm requires some collective, common, or mutual cognitive states the contents of which reciprocally refer to the contents of others' cognitive states about others' cognitive states. This approach is quite common in analytical philosophy.[7]

The model is clearly cognitive. It first requires of an individual that she has conceptual knowledge of mental states and an understanding of under what conditions they apply to others, and that she believes that other have this understanding as well. Secondly, it requires that her understanding be rationally inferred from interpretations of mental states of others whom she must believe to engage in the same activity of rational inference and interpretation. This means that a norm and the attainment of an understanding of the norm both require that all individuals whom are subject to it have a folk psychological theory by means of which they infer and attribute the relevant preferences, beliefs, and expectations.

For a second illustration of a cognitive model we may consider Cristina Bicchieri's, presented in *The Grammar of Society* (2006). Bicchieri defines a social norm (p. 11) as follows. It is a behavioural rule R, for situations of type S in a population P such that: if there exists a sufficiently large subset of conditional followers $P_{cf} \subseteq P$ such that, for each individual $i \in$ [member of] P_{cf} then,

> *Contingency*: i knows that a rule R exists and applies to situations of type S;
> *Conditional preference*: i prefers to conform to R in situations of type S on the condition that:
> (a) *Empirical expectations*: i believes that a sufficiently large subset of P conforms to R in situations of type S;
> and either
> (b) *Normative expectations*: i believes that a sufficiently large subset of P expects i to conform to R in situations of type S;
> or
> (b´) *Normative expectations with sanctions*: i believes that a sufficiently large subset of P expects i to conform to R in situations of type S, prefers i to conform, and may sanction behavior.

That a social norm R is *followed* in a population P, on Bicchieri's definition (ibid), means that:

> there exists a sufficiently large subset $P_f \subseteq P_{cf}$ such that, for each individual $i \in P_f$, conditions 2(a) and either 2(b) or 2(b´) are met for i and, as a result, i prefers to conform to R in situations of type S.

It is sufficient for categorizing this as a cognitive model of acting in accord to norms, according to my scheme of categorization, that Bicchieri claims that for an individual to follow a norm she must minimally have certain true beliefs (knowledge) about others' regular behaviour in specific situations and prefer to conform to the regularity in the relevant type of situations, conditional on her having these beliefs about others' behaviour in the relevant type of situation and about their expectations about her behaviour (or about their expectations and preferences about her behaviour and their dispositions to sanction her). That this qualifies as a cognitive model should be obvious once it is considered what is required of any one individual in order that we may justifiably attribute to her an understanding of social norms: she must, first, have a folk psychological theory about mental states and, secondly, infer what to do from her knowledge that others instantiate certain mental states, as conceptualized in that theory, or from their mental states and dispositions to sanction deviation from behavioural regularities in certain situations.

It should be noted that the relevant cognitive processes that cognitive models tend to posit might be implicit and not consciously or explicitly performed or monitored by individuals to whom they are attributed. That is, the fact that meta-levels of knowledge or belief are required for understanding norms need not be a fact about at what level of cognition, conscious or non-conscious, the processes occur. One might not, for example, be required to have higher-order beliefs that one have beliefs about or performs the relevant inferring and attributing of mental states. It is sufficient for the model that the posited processes of formation, inference, and attribution of mental states occur.

The cognitive and non-cognitive models, as here presented, should provide some clarification of the wider field of research that this dissertation relates and contributes to, as well as specification of the meaning of terms as they occur in the papers that follow. In the next section I clarify in

more detail how I conduct this project of contributing to the cognitive–non-cognitive debate.

2.4 Method and Hypothesis

I assess the cognitive and non-cognitive models to social understanding and normativity by means of a stepwise, paper by paper, investigation of one version of the non-cognitive model, either in itself or as pitched against a cognitive model. In this manner I assess, in the first four papers, the prospects for a situated, embodied, enactive, and ecological approach to how people understand each other and how they understand and act according to norms.

The method in the relevant papers is theory development and argumentation analysis. The first represents my endeavour to, to begin with, formulate the respective approaches to norms and social understanding, and how we act accordingly. This also involves contrastive analysis. By this is meant that in the investigation of a non-cognitive approach to, e.g., social understanding or joint action according to norms, I present a proposed cognitive model and then develop a non-cognitive model as a counterweight or contrast to it.

The second part, the argumentation analysis, consists in assessing the strengths and weaknesses of arguments for a non-cognitive model in the relevant domain, with an aim to specify how to develop it as an improvement on a cognitive model. A problem in this connection is how to understand what saying that one model 'improves' on another means. I mean by it that the model developed has at least one of the following properties:

1) *Generality:* It presents a more *general* understanding than the competitor model. For example, it finds gaps, or counterexamples, to the competitor and shows how these can be handled. A model that helps us both see gaps in generality and how to fill these gaps is to be regarded as more of an improvement than one that only does the first.

2) *Explanatory power:* It *better explains* a range of phenomena than the competitor model. For example, it explains why on the competitor a certain instance of the phenomenon investigated presents a problem for the competitor and how this instance can be explained in the model developed. A more challenging model that also handles the challenge without loss in generality is to be regarded to be more of an improvement than one that points to many challenges but cannot handle them itself.

That a model is more general or explanatorily more powerful than another can have either of two consequences. First, it might be that aspects of the challenging model should be *added* to the model challenged. In this case the latter can be retained but stand in need of auxiliary claims from the former. The challenger points to something that the challenged model cannot handle, a counterexample, but that can be handled by adding some of the challenger's claims. Secondly, it might be that certain claims of the challenged model must be *replaced* (or abandoned). In this case it cannot be retained as it stands. The challenger points out that the challenged model gives a wrong verdict for some range of cases. Replacement may be suggested in the challenging model, in which case the challenged can either accept the suggested replacement or retract and thus loose in scope of generality. Other properties of a model that suggests that it is an improvement on a competitor are:

3) *T-Simplicity:* It is *theoretically simpler* than the competitor. By this is meant that the number of claims and theses the model must muster in order to make sense of a phenomenon are fewer than on the competitor. That is, the model lets us understand more with less theory.

4) *M-Simplicity:* It is *metaphysically simpler* than the competitor. By this I mean that the model leaves what is understood simpler by representting the world as containing fewer kinds of entities.

5) *Coherence:* Its claims and theses, or those that can be derived from it, are mutually supporting, or at least not mutually undermining in the sense that accepting one excludes accepting another or that the consequences of one implies the falsity of (the consequences of) another.

I do not arrange these virtues in any order of importance. Also, several more specifications of 'improvement' could be added, but I have wanted to keep the list short and T-simple. Moreover, I include only those that will turn out relevant for the papers that follow. Nevertheless, something should be said, at least by way of example, of how the five are related.

Suppose, for instance, that model **M1** provides a general explanation of action in accord to norms and how people understand norms by means of employing only a few concepts such as embodiment, enaction, behavioural pattern, and affect, where each of these has a conservative extension such that it requires only, say, four kinds of processes in the world. However, as it turns out, a consequence of the meaning of 'affect', as understood in **M1**, implies a contradiction because of some claim made regarding the meaning of 'behavioural pattern'. On the other hand we have model **M2**, which employs fever concepts, e.g., common knowledge, preference, and behavioural pattern, which are fully compatible but which suffers in generality, as shown in several counterexamples produced by **M1**, and which also populates the world with a whole range of kinds of mental processes. Suppose that **M1** and **M2** are incompatible. The question for us is which of **M1** or **M2** to accept, if either. I do not propose an answer here. In interesting borderline cases like these we may have to await further investigation. But if one model is clearly an improvement on another because of the listed properties it exhibits, then this may guide our judgment.

The problem for the method of theory development and argumentation analysis regarding a metric for assessing the relative improvement of one model on another should be ameliorated with the help of the provided list. I have wanted to make the list as theory-neutral as possible—not with respect to theories in the philosophy of science but with respect to the cognitive and non-cognitive models that are my concern. The list does not, as far as I can see, beg any important questions against either model.

A problem with the list is that a class of properties that might be thought to be of chief importance for assessing a model is omitted. We may call it the class of veritistic properties; that the model truly or with some sufficiently high probability represents the reality of what is modelled, The claim that veritistic properties are important is plausible. I am not much

concerned with the truth-value of claims, however, but with the relative superiority of models at a stage where it is not clear what their veritistic properties are. And even if one turns out to be less probably a true representation than a competitor it might still be true that it is conceptually more persuasive than the competitor—according to the submitted list. Supposing that model **M1** is less probably a true representation of what is modelled than competitor model **M2** but that **M1** satisfies the other conditions of improvement, it is a further question whether the predicament is culpa the model or culpa the (interpretation of) observations or experiments that supposedly make it less probably true. I do not make such suppositions and I do not engage the question what to do in the hypothetical circumstance that such suppositions were true.

The hypothesis that this dissertation investigates is: a non-cognitive model of the nature of human understanding of norms and of others, and what is involved in acting accordingly, is on the whole an improvement on a cognitive model.

However, another issue that will be tackled, or actually two issues, is this: is meaning and belief essentially normative, in the sense that we cannot understand a word or expression as having meaning or a mental state as a belief without understanding them as governed by norms? The reason why I have devote this question a fair amount of attention is that it seems to strike at the heart of what is often taken to be distinctly human activities. While all animals interact with each other in a variety of manners it seems that only humans form, revise, criticize, and question what they believe. And only humans seem to engage in the complex form of communication and language use that comes with having our capacity to attach meaning to noises or material symbols (like we do with this clause), and to discimriatne crorect fmor inorcerct seu (like we do with this clause). Hence I have found it important to discuss these two phenomena, meaning and belief, separately, because they tend to be characterized as human-specific, or at least specific to species that act for reasons.

In the context of meaning and belief I go about the issue whether they are essentially normative with a method similar to that for the first four papers: argumentation analysis. That is, I consider the arguments for and against the relevant claims that meaning and belief is normative and assess their strengths and weaknesses, not relative to the list above but relative to

their validity. Specifically, the hypotheses investigated in the last two papers are: first, that meaning, in the sense that the fact that a word or expression has a certain semantics, entails norms governing its use; and, secondly, that a certain class of beliefs lend support to the general claim that the fact that something is a belief entails that norms govern whether to accept, revise, or reject it.

Notes:

[1] The distinction between cognitive and non-cognitive models is formulated in the description of the project *Understanding the Normative Dimensions of Human Conduct: Conceptual and Developmental Issues* (grant no. 429-2010-7181) to which this dissertation belongs. The specification of the two models in this chapter is thus also a specification of an assumption made in that project, with an aim to investigate the two models' prospects for accounting for how people understand and act according to norms. An issue that I will not tackle is whether the two models can be made compatible.

[2] Here I might add that others have counted to more than three senses of the term 'embodied' and also that others give a somewhat different characterization than I give (see, e.g., Wilson 2002; Anderson 2003; Ziemke 2003). In my count and definitions I follow Kiverstein (2012) and Clark (2008a).

[3] I am in this dissertation taking an offense-position rather than a defense-position with respect to embodiment. That is, when I investigate a (strong) embodiment approach to social understanding and our understanding of norms I am mainly concerned with how such an approach is to be formulated, rather than with how it is to be defended given the proposed formulation(s). This is true also in the context of my formulation of the ecological approach to normativity. For this reason the reader may want at least to be given directions to objections that may be raised. Such can be found in, e.g., Adams and Aizawa (2010a; 2010b), Adams (2010), and Aizawa (2014). Also, for debate between different versions of non-cognitive models: see, e.g., Hutto (2005), Clark (2008b), Kiverstein and Clark (2009), Gallagher and Bower (2014), and Stapleton (2013). In the case of the situated approach I do consider and respond to several objections.

[4] What I call 'situated' here can, without much change having to be made, be likened to what has been proposed under a different label; namely 'embedded'

understanding or cognition. See, e.g., Haugeland (1998), Clark (2008b), Crisafi and Gallagher (2010), and Gallagher (2013).

5 Here I might add that there are also several suggestions how to interpret the term 'affordance'; for example whether an affordance is a property *of the environment*, which suggests that affordances have objective reality in abstraction from a kind of animal. In this sense the affordances of a human niche, for example, would continue to exist despite all humans going extinct because the environment would still afford humans certain actions. For more discussion on how to interpret 'affordance' see, e.g., Greeno (1994), Costall (1995; 2012), Chemero (2003; 2009), Heft (2003; 2013), Heras-Escribano and de Pinedo (2015), and Kiverstein (2015).

6 It should at least be mentioned here that recently a *radical* sense of enactivism has been proposed (Hutto and Myin 2013; Myin and Hutto 2015). I have not yet had the opportunity to investigate what consequences a radicalization of enactivism has for the relevant enactivism-inspired papers in this dissertation.

7 See, for example: Gilbert (1989; 2013), Tuomela (2007), and List and Pettit (2011).

3. Summary of the Papers

I will now summarize the main arguments and conclusion of the papers taken separately. In the next part I will draw the common conclusions to which they point.

Paper 1 (*Situating Norms and Jointness of Social Interaction*, Lo Presti 2013a) deals with the situated approach to understanding norms and how people can be understood to act jointly in accord to norms.

The hypothesis that I attempt to vindicate in this paper is that understanding how to act jointly with others according to social norms does not generally require having a folk psychological theory about how to 'read the mind of others'. Such a theory involves having mental state concepts, such as 'belief', 'preference', and 'expectation', which individuals must know the conditions for attributing to others, where such attribution is premised on an inference from observing how others behave in the context of a particular situation. Thus the argument is that a cognitive model approach to joint action in accord to social norms must be somewhat restricted in its scope because there are counterexamples to it, which, I argue, can be handled on an alternative non-cognitive model.

With respect to the five properties that an alternative model must exhibit in order to qualify as an improvement on a competitor, the situated approach is argued to support the claim that a cognitive model suffers in the arena of generality because counterexamples can be produced. The situated approach is argued to be able to handle these counterexamples. Hence this paper establishes that a situated version of the non-cognitive model to joint action in accord to norms must in many cases be added to

our understanding of this phenomenon by way of replacing certain theses of a more traditional cognitive model. And the way that this claim is supported is by means of argumentation analysis and theory development. That is, versions of the cognitive model are first considered, counterexamples are presented, and then the situated approach is developed and argued to handle the counterexamples. Ergo, in many cases the cognitive model fails, and in these cases a situated approach must be appealed to in its stead.

The first paper, then, throws us directly into the debate between cognitive and non-cognitive models without considering in detail what a norm is supposed to be apart from some standard of correctness or appropriateness. It concludes that understanding our understanding of norms and how to act jointly as a situated understanding is called for with respect to many everyday situations because otherwise our understanding of that understanding will lack generality. This strategy, as clarified in the previous chapter, is symptomatic of the first four papers of this dissertation.

Paper 2 (*Social Ontology and Social Cognition*, Lo Presti 2013b) introduces an enactivist, non-cognitive, approach to the construction of social environments and of social understanding as a counterweight to an orthodox cognitive model.

The cognitive model considered here is that social environments require, for their existence, that people together explicitly declare and impose norms or rules on their interactions, or have common knowledge or mutual beliefs that such apply. The enactive approach, on the other hand, understands the creation of social environments as a process in which people enact roles in social interaction, where this does not require previous explicit declaration or imposition of roles or statuses.

The process of enaction of social roles and statutes is here characterized as an ongoing social interaction that help people perceive what they do and their environment as having a socially significant structure. Because of this, participation in social interaction in the context of a particular situation is participation both in the enaction of social roles and in a context in which social roles have been established over time. This means that participatory understanding can be described as creating a social reality—social roles, statuses, groups, institutions—and that

participating in social interaction in the social reality thus created relative to a particular group allows one to understand others and how to act in its context.

So the hypothesis that I attempt to vindicate in the second paper is that an enactive approach to the creation of social reality and to people's understanding each other and how to interact, as agents situated together in a social reality, is a viable alternative to a cognitive model of these same phenomena. And the strategy here is again to contrast, first, the cognitive and enactive models as competitors and, secondly, to develop the enactive model and to point to its virtues.

This paper also contains a review of the history of ideas regarding the creation of social reality. Here the contemporary analytical tradition in philosophy is seen to be just one cognitive approach that has rapidly become orthodoxy despite a long tradition in sociology that seems rather to support a non-cognitive model. It is with one leg on each of the shoulders of these two traditions that the second paper takes a step towards a non-cognitive, enactive model on social reality and social understanding. The conclusion is twofold. First, since research on the creation and structure of social reality tends to agree that these are to be understood in terms of people's social understanding, research on social reality must be sensitive to results in research on social understanding, and vice versa. Secondly, since, according to my argument, social reality can be understood as enacted and social understanding as enacting a social reality, social understanding and social reality must be understood as involved in a relation of co-dependence, in the sense that how people understand each other and their environment in social interaction tend to produce social significance, which tend in turn reciprocally to produce social understanding (or misunderstanding).

The first two papers taken together provide a situated-enactive account of the structure of social reality and how it is that people learn, understand, and act jointly according to norms. They also show how people can understand each other as subjects situated together in the concrete context of an enacted social environment in particular situations. At this point proponents of a non-cognitive model should begin to look not as radicals without a history or common cause, scattered castaways on the fringes of what requires attention, but on the contrary as a joint force with its own coherent story to be told and reckoned with. Reckoning is what I

do and where doing so leads us is what I investigate for those not yet aware.

Paper 3 (*Rule-Following, Meaning Constitution, and Enaction*, Lo Presti 2015) examines whether an enactivist solution to Wittgenstein's rule-following paradox, as interpreted by Saul Kripke, is available. This is contrasted to sceptical and straight solutions, which have traditionally been supposed to mutually exhaust possible solutions. The sceptic accepts the paradox—no fact constitutes our meaning something by an expression—whereas the 'straight' insists that it can be circumvented—there is some fact constitutive of the meaning of an expression. The straight faces a dilemma. On the one hand we have the straits of Scylla: on this monster's horns we have to concede that no rule that we have in mind determines what an expression means such that there is a specific way how to use it, because use exemplifies an indefinite number of rules, none of which can be singled at as the correct rule under all circumstances. On the other side we have the straits of Charybdis: on this monster's horns we have to admit that mere regularities in linguistic behaviour, what we are disposed to do with an expression, is insufficient for determining what an expression means in all cases, because we are disposed to many things that do not seem to constitute what we mean with an expression.[1]

The hypothesis here is that the gulf between meaning being constituted either in mental representations of a rule as exemplified in a practice, or in a behavioural disposition to pursue a practice in some regular way, both have a common ground and can be traversed. Specifically, it is argued that both Scylla and Charybdis are figments of a common imagination or orthodoxy that posits an essential gulf between mind and behaviour where instead it might be claimed that mind is embodied and that embodied action enacts meaning on a case-by-case basis as we go along. On the Scylla and Charybdis metaphor, the enactivist solution that this paper investigates equips us with sonar that reveals that the same seabed below connects the two sides. When we understand this we may also chart our course straight, arguing that meaning is not constituted either in bodily behaviour or mental representation but in embodied action, understood as constitutive of our understanding what to do. Less metaphorically, the idea is that the meaning of an expression or action

consists not *either* in internal, disembodied, mental representations of rules for how to go on *or* in external, mindless, behavioural patterns, but is enacted in a processes of embodied social interaction.

What the paper shows is only *that* enactivism can be formulated so as to read the rule-following paradox in a new, non-cognitive, light, not *how*, precisely, it should be formulated. The paper leaves us with the conclusion that the either-or story of old, which keeps feeding the Sicilian and Calabrian monster-myths, clouds our minds from a sober re-assessment: the dilemma is two heads of the same myth of mind as disembodied and embodiment as mindless. What to do with this conclusion is left for another occasion.

Together, the first three papers challenges the cognitive model: first, on what is required for people to understand how to act jointly in accord to social norms; secondly, on what is required for people to understand each other in a social context that they participate in enacting, as well as on what this participation and enaction consists in; and, finally, on what constitutes meaning, whether of linguistic practice specifically or social practice more generally. In each case the challenger model is presented by means of contrastive analysis of the arguments for and against the two sides, followed by a development of a situational, enactive, and embodied non-cognitive model, argued to be an improvement to the challenged cognitive model.

Paper 4 (*An Ecological Approach to Normativity*, Lo Presti In Press) returns to and deepens the investigation of a non-cognitive model for how people learn, understand, and act according to norms.

The hypothesis is that norms can be understood as affordances of the ecological niche of a social kind of animal such that its inhabitants are directly presented not just with *possibilities* for action but also, among possible actions, with how to discriminate those that are normative from those that are not. This direct understanding of norms is understood as an embodied practical know-how enabled by way of participation in interaction in the pursuit of social practices in the relevant niche. Such participation also tends to produce and change norms of the niche. Therefore also the establishing of norms is, itself, understood as a dynamic process in the concrete context of social interaction. This is contrasted to a cognitive model approach on which norms are realized by way of people stating that

norms apply, as expressed in declarative speech acts, and on which learning and understanding what norms apply is exhibited in one's capacity to represent them in a propositional form stating *that* something is a norm. This, it is argued, is neither necessary nor sufficient for something to be correct, nor for people to learn how to act correctly. One may know *that* something is correct in the context of some practice without being able to handle actual engagement in the practice correctly, and one may know *how* to act correctly in the context of some practice without knowing or being able also to say *that* it is correct, or even being able to say that one know how to act.

If norms can plausibly be considered to be affordances of a niche of social practices then norms must not be assumed to consist either in the bodily behaviour of individuals or in disembodied representations of some behavioural regularity as a norm, which is how the cognitive model has us picture it. As is true of the hypothesis in the previous paper with regard to meaning constitution being a dynamic enactive process, i.e., a process in which embodied activity continuously produces and changes meaning, also norms, in this paper, must be understood as dynamically enacted, i.e., as constantly established and changed or overthrown in the concrete context of embodied social interaction in a niche of social practices.[2] This is a non-cognitive, ecological, contrast to a cognitive approach to what kind of fact must be supposed to obtain for it to be true to say that something is a norm and that people understand this and act accordingly. The kind of fact in question, on the hypothesis examined, is essentially relational because ecological, and relational facts are essentially dynamic in the sense that they constantly change as a consequence of what people do in particular situations over time.

The strategy adopted is thus theory development against a background of contrastive argumentation analysis, where a cognitive model is considered and pitched against a non-cognitive, this time ecological, challenger model. Here the improvement of the challenger is foregrounded against a background of dilemmas argued to beset the view that norms and activity in accord to norms is akin to rules and rule-following behaviour. The ecological challenger is argued to throw new light on Wittgenstein's poignant but obscure notions that rule-following must ultimately be "blind", i.e. without justifications, at the level of "bedrock" activity, i.e. at the level

where we cannot give a further interpretation of what we do as being in accord to a rule and have to say "This is simply what I do".

The first four papers constitute the advancement of a non-cognitive model on how normativity and meaning is constituted, understood, and acted in accord to, but also on how people understand each other. The remaining two papers ask whether meaning is normative and whether belief is normative.

Paper 5 (*Speaking About the Normativity of Meaning*, submitted) engages the debate between what has become known as 'normativism' and 'anti-normativism' about meaning. According to normativism, meaning is normative in the sense that a word or expression can be said to have meaning only if it is subjected to correctness conditions, i.e. only if under certain circumstances it is correctly used and, in others, incorrectly. In contrast, random marks or noises that are not so distinguishable into those correctly used and those incorrectly used are said to have no meaning. Anti-normativist, in contrast, analyse 'correctness' as a descriptive term referring to whether a word or expression has some non-normative property, e.g. that in using it one is saying something that is true or false, or, indeed, that in saying that it is used correctly or incorrectly one is only describing it as conforming to some standard of semantic categorization by means of noise-making that is not itself normative.

The hypothesis examined in this paper is that meaning is not normative in the normativist sense, and yet meaning is not non-normative. The idea is that while meaning facts, i.e. that a word or expression has some specific meaning, does not entail any norm saying that there is some way in which one ought or ought not use it (which is what normativists claim), still the fact that a word or expression has meaning to begin with presupposes certain normative properties in patterns of use. These are properties of: committing and entitling a speaker to certain words or expressions given words and expressions used; committing and entitling a hearer to ask the speaker for correction or reasons in case the hearer does not accept what the speaker says; entitling the hearer to defer to the speaker and commit to what the speaker says if the hearer accepts what the speaker says. This is not the idea that people ought or ought not speak and use words in this way or that but that unless they commit and entitle in the

sense specified then this imperils their status as speakers and hence as someone to be listened to. On this view one would not violate any norm essential to meaning if one, e.g., speaks falsely, because meaning does not entail any such norm. But, on the other hand, one can be recognized as meaning something by a noise or symbol if and only if one can be counted on to commit and entitle others to commit to certain further expressions.

The conclusion here is that anti-normativism can be accepted without this entailing that one accepts that meaning is non-normative, because the normativism that one would thereby be rejecting does not cover the whole field of meaning of the claim that meaning is normative. Meaning can still be normative in the sense that meaning facts presuppose normative properties of patterns of making noise and marks even if the fact that those noises or marks have meaning does not in turn entail that one ought or ought not use them in a particular manner.

Paper 6 (*Moore's Paradox and Epistemic Norms*, Lo Presti 2014) investigates another normativist--anti-normativist debate. This time the debate pertains to whether a solution to Moore's paradox, at least in one of its various formulations, can be given in terms of the alleged normativity of belief. The hypothesis here is that this class of beliefs must be understood as essentially normative in order that we can understand why they are paradoxical.

The analysis of arguments on both sides that I present in this paper contradicts the hypothesis. Normativism, i.e. the position that the reason why certain beliefs strike us precisely as *paradoxical* is because the believer in having them violates norms, is not plausible. The reason why is that the normativist explanation, properly analysed, requires that for someone to be said to violate the relevant norm he or she must, minimally, believe that this is the case. Hence, in order to violate the norm the individual must have a belief that her belief violates the norm. But if belief is normative then the belief that the belief violates the norm is also subjected to the norm of belief. This means that the believer, in order properly to violate the norm and hence to be said to have paradoxical beliefs, must have a belief that the belief that violates the norm violates the norm, and he or she must of course also have a belief that the belief that the belief... that violates the norm

violates norm, and so on *ad infinitum*. The normativist explanation of Moore's paradox, it is argued, is itself absurd.

Another normativist explanation that is considered is that one *cannot really believe* both the propositions that in conjunction constitute a Moore-paradoxical belief ('*P* and I do not believe that *P*', or '*P* and I believe that *not-P*'), because if one is true then the other must be false. On this proposal, the claim that belief is governed by a truth-norm that one ought to believe only the truth is to be understood as a conceptual truth. Hence one *cannot, by definition*, really have Moore-paradoxical beliefs because they contain the beliefs that something is the case and the belief that one does not believe that it is the case (or the belief that one believes that it is not the case). It is argued in the paper that this explanation is self-defeating. It cannot explain Moore-paradoxical beliefs. The reason is that if we are supposed to understand a belief as paradoxical because one *cannot as a conceptual truth have it*, then it is difficult to make it intelligible how there is something to explain in the first place. The explanation takes the general form: *X* is paradoxical because *X* is impossible. But, now, if *X* is impossible then how can this explanation explain the occurrence of *X* to begin with? And yet Moore-paradoxical beliefs, this normativist proposal has it, are to be explained by reference to the normativity of belief. However, Moore-paradoxical beliefs cannot be so explained since the explanation given effectively erases the possibility of what was to be explained; i.e., it says that you cannot have the beliefs. If it does not render the explanandum inexplicable, then the normativist theory from which the explanation is derived must be wrong. This is so because if a Moore-paradoxical belief is to be understood as an instance of a violation of a norm which, according to the theory, one cannot really violate, then it cannot be the case that one cannot really violate the norm. Therefore, if the explanation is correct then the normative theory that supports it must be false, and if the normative theory is true then the explanation must be incorrect.

Taken together, these six papers is an investigation into whether a non-cognitive model how norms are established, understood and acted in accord to, and how people understand each other and how to act jointly, is an improvement on a cognitive model. The papers also provide an

assessment of the arguments that meaning and (a special class of) belief are essentially normative.

Notes:

[1] The terminology and reference to the Homeric episode where Odysseus has to steer his ship straight between the sea-monsters Scylla (Sicily) and Charybdis (mainland Italy), representing the two horns of a dilemma, is borrowed from McDowell (1984).

[2] The similarity to the previous paper is brought out more clearly if we understand the cognitive model approach as follows. The fact of the matter whether something is a norm and whether people understand this and act accordingly is a fact *either* about their beliefs, expectations, preferences, and inferences about each other's beliefs, expectations, preferences, and inferences about each other's beliefs, expectations, preferences, and inferences about each other's... and so on, *or* about patterns in behaviour conceived as of no consequence for people's understanding of each other and of norm s. Thus understood, the ecological approach, on which the fact of the matter rather is about relations spanning embodied social interaction in a niche of affordances, is an approach that points out, once again, that the Scylla (mentalist) and Charybdis (behaviourist) straits are myths of monsters fed by the notion that understanding is essentially disembodied and embodied activity of no consequence for understanding.

4. Conclusions

The general conclusion of this dissertation is that a non-cognitive model of social understanding, understanding of norms and how to act correctly, joint action, and the constitution of meaning and construction of social reality, is a strong competitor to cognitive orthodoxy. The non-cognitive model produces counterexamples that threaten the generality of the cognitive model and develops an alternative that helps understand the problem cases. The non-cognitive model shows that, often, complex conceptual capacities required on the cognitive model, like reading other minds by means of inferring and attributing mental states from observations of behaviour, and having propositional knowledge of rules that represent or impose norms on conduct, are not necessary for understanding others, the environment, and how to act according to norms in the context of many particular situations.

When the embodied, situated, enactive, and ecological nature of human understanding of norms is properly understood a non-cognitive model sails up as a cohesive and fast moving challenger. It challenges not only orthodox cognitive science, in which the cognitive model has its home port, but long-lived beliefs about philosophical puzzles, such as: that mind and body, and body and world, are to be understood as fundamentally different—that we can, perhaps even should, understand how and what we think in abstraction from it being an activity involving what we do with our bodies in interaction with others in particular situations; that understanding other minds is a matter of theorizing from inference and analogy to one's own mind from observed behaviour that supposedly hides the other's mental states; that meaning is determined by rules interpreted as

applying in an indefinite number of cases; that social reality exists only in the mind of a collective by way of its members rationally believing that they collectively believe that they collectively believe that they collectively believe that... they collectively believe that it exists. I conclude that the manner in which the non-cognitive model here examined calls these claims into question ought, at the very least, to be reckoned and answered to by anyone who makes such claims as if they needed no support. If anything, I hope that this dissertation contributes to lifting the fog between proponents to the two models so that they may better understand each other's claims and hold each other committed and entitled to claims that they are entitled to demand reasons for and committed to give reasons for.

Another thing that should not be forgotten is the conclusion that as the debate between normativists and anti-normativists about meaning is now raging there seems to be another sense in which meaning is essentially normative which is not *their* difference of opinion. This is the sense in which meaning is normative because there is no meaning or content to words or thoughts in abstraction from uses of noises or marks (like these written here) having the properties of committing and entitling to words and thoughts. Now, once meaning is established and associated with some word or thought it may be expressed and thought in whatever manner one sees fit without violating any norm essential to that word or thought. But this freedom is possible only because the word or thought is understood as having meaning, which presupposes that it is recognized that using the word or thinking the thought commits and entitles to demand responsibility for it and to take responsibility for it if this is demanded. This is an ongoing contest over what norms should be accepted as applying and when (e.g. a truth-norm in the context of public debates). These are *contingent* norms. The *essential* norm, which applies in *all* circumstances, is to commit and entitle, at least implicitly, i.e. whether or not one can say that one does so. If one can speak or think about commitments and entitlements that are at least implicit in use then one can also speak and think meaningfully about the normativity of meaning.

Finally, it is concluded that belief, at least in the context of Moore-paradoxical beliefs, does not entail norms. It is not required to understand why these beliefs should strike us as paradoxical that they constitute

violations of norms that apply to all beliefs. Indeed, the idea that that is how we should understand the paradox is itself tied up with absurdity.

The three conclusions, succinctly put, are: a non-cognitive model of the nature of human understanding of norms and others, of creating social environments and of understanding how to act jointly according to norms, is at the very least a plausible challenger to a cognitive model; meaning can be understood as essentially normative without this implying that any particular ought or ought not is essential to meaning; and, finally, Moore's paradox is not to be understood in terms of violations of norms that apply to all beliefs as a matter of conceptual truth.

References

Adams, F., and Aizawa, K. (2010a). *The Bounds of Cognition*. Malden, MA: Blackwell.

Adams, F., and Aizawa, K. (2010b). 'The value of cognitivism in thinking about extended cognition'. *Phenomenology and the Cognitive Sciences* 9, 579–603.

Adams, F. (2010). 'Embodied cognition'. *Phenomenology and the Cognitive Sciences* 9, 619–628.

Aizawa, K. (2014). 'What is this cognition that is supposed to be embodied?' *Philosophical Psychology*, Online First: DOI 10.1080/09515089.2013.875280.

Anderson, M. L. (2003). 'Embodied cognition: a field guide'. *Artificial Intelligence* 149, 91–130.

Bicchieri, C. (2006). *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York, NY: Cambridge University Press.

Brandom, R. B. (1994). *Making it Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.

Brandom, R. B. (2000). *Articulating Reasons: An Introduction to Inferentialism*. Cambridge, MA: Harvard University Press.

Brandom, R. B. (2008). *Between Saying and Doing: Towards an Analytic Pragmatism*. New York, NY: Oxford University Press.

Brinck, I. (2015). 'Understanding social norms and constitutive rules: perspectives from developmental psychology and philosophy'. *Phenomenology and the Cognitive Sciences*, Online First: DOI 10.1007/s11097-015-9426-y.

Chemero, A. (2003). 'An outline of a theory of affordances'. *Ecological Psychology* 15(2), 181–195.
Chemero, A. (2009). *Radical Embodied Cognitive Science*. Cambridge, MA: The MIT Press.
Clark, A. (1999). 'An embodied cognitive science?' *Trends in Cognitive Sciences* 3(9), 345–351.
Clark, A. (2008a). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. New York, NY: Oxford University Press.
Clark, A. (2008b). 'Pressing the flesh: a tension in the study of the embodied, embedded mind?' *Philosophy and Phenomenological Research* 76(1), 37–59.
Clark, A., and Chalmers, D. (1998). 'The extended mind'. *Analysis* 58(1), 7–19.
Costall, A. (1995). 'Socializing affordances'. *Theory & Psychology* 5(4), 467–481.
Costall, A. (2012). 'Canonical affordances in context'. *Avant* 3(2), 85–93.
Crisafi, A., and Gallagher, S. (2010). 'Hegel and the extended mind'. *AI & Society* 25, 123–129.
De Jaegher, H., and Di Paolo, E. A. (2007). 'Participatory sense-making: an enactive approach to social cognition'. *Phenomenology and the Cognitive Sciences* 6, 485–507.
De Jaegher, H., and Di Paolo, E. A. (2013). 'Enactivism is not interactionism'. *Frontiers in Human Neuroscience* 6, 1–2.
De Jaegher, H., Di Paolo, E. A., and Gallagher, S. (2010). 'Can social interaction constitute social cognition?' *Trends in Cognitive Science* 14(10), 441–447.
Dreyfus, H. L. (2007). 'The return of the myth of the mental'. *Inquiry* 50(4), 352–365.
Dummett, M. (1973). *Frege: Philosophy of Language*. New York, NY: Harper and Row.
Engel, P. (2013). 'Sosa on the normativity of belief'. *Philosophical Studies* 166, 617–624.
Froese, T., and Gallagher, S. (2012). 'Getting interaction theory (IT) together'. *Interaction Studies* 13(3), 436–468.

Fuchs, T., and De Jaegher, H. (2009). 'Enactive intersubjectivity: participatory sense-making and mutual incorporation'. *Phenomenology and the Cognitive Sciences* 8, 465–486.

Gallagher, S. (2004). 'Situational understanding: a Gurwitschian critique of theory of mind'. In L. Embree (Ed.) *Gurwitsch's Relevancy for Cognitive Science* (25–44) Dodrecht: Springer.

Gallagher, S. (2013). 'The socially extended mind'. *Cognitive Systems Research* 25-26, 4–12.

Gallagher, S., and Bower, M. (2014). 'Making enactivism even more embodied'. *Avant* 5(2), 232–247.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. New York, NY: Psychology Press.

Gibbard, A. (2014). *Meaning and Normativity*. New York, NY: Oxford University Press.

Gilbert, M. (1989). *On Social Facts*. Princeton, NJ: Princeton University Press.

Gilbert, M. (2003). 'The structure of the social atom: joint commitment as the foundation of human social behaviour'. In F. F. Schmitt (Ed.), *Socializing Metaphysics: The Nature of Social Reality* (39–64). Lanham, MA: Rowman and Littlefield.

Gilbert, M. (2013). *Joint Commitment: How We Make the Social World*. New York, NY: Oxford University Press.

Glüer, K., and Wikforss, Å. (2009). 'Against content normativity'. *Mind* 118(469), 31–70.

Greeno, J. G. (1994). 'Gibson's affordances'. *Psychological Review* 101(2), 336–342.

Grice, P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Gurwitsch, A. (1931/1978). *Human Encounters in the Social World*. Ed. by A. Métraux, trans. by F. Kersten. Pittsburgh, PA: Duquesne University Press.

Haugeland, J. (1998). *Having Thought: Essays in the Metaphysics of Mind*. Cambridge, MA: Harvard University Press.

Hattiangadi, A. (2007). *Oughts and Thoughts*. New York, NY: Oxford University Press.

Heft, H. (2003). 'Affordances, dynamic experience, and the challenge of reification'. *Ecological Psychology* 15(2), 149–180.

Heft, H. (2013). 'An ecological approach to psychology'. *Review of General Psychology*, 17(2), 162–167.

Heras-Escribano, M., and de Pinedo, M. (2015). 'Are affordances normative?' *Phenomenology and the Cognitive Sciences*, Online First: DOI 10.1007/s11097-015-9440-0.

Hutto, D. D. (2005). 'Knowing *what*? Radical versus conservative enactivism'. *Phenomenology and the Cognitive Sciences* 4, 389–405.

Hutto, D. D., and Myin, E. (2013). *Radicalizing Enactivism: Basic Minds without Content*. Cambridge, MA: The MIT Press.

Jackson, F. (1999). 'Non-cognitivism, normativity, belief'. *Ratio* 12, 420–435.

Kiverstein, J. (2012). 'The meaning of embodiment'. *Topics in Cognitive Science* 4, 740–758.

Kiverstein, J. (2015). 'Empathy and the responsiveness to social affordances'. *Consciousness and Cognition*, Online First: DOI 10.1016/j.concog.2015.05.002.

Kiverstein, J., and Clark, A. (2009). 'Introduction: mind embodied, embedded, enacted: one church or manu?' *Topoi* 28, 1–7.

Kripke, S. (1982). *Wittgenstein on Rules and Private Language*. Malden, MA: Blackwell Publishing.

Lewis, D. (1969). *Convention*. Malden, MA: Harvard University Press.

List, C., and Pettit, P. (2011). *Group Agency*. New York, NY: Oxford University Press.

Maturana, H. R., and Varela, F. J. (1972/1980). *Autopoiesis and Cognition: The Realization of the Living*. Dodrecht: D. Reidel.

McDowell, J. (1984). 'Wittgenstein on following a rule'. *Synthese* 58, 325–363.

McGinn, C. (1984). *Wittgenstein on Meaning*. Oxford: Blackwell.

McGinn, C. (2011). *Truth by Analysis: Games, Names, and Philosophy*. New York, NY: Oxford University Press.

Moore, G. E. (1942). 'A reply to my critics'. In P. A. Schlipp (Ed.), *The Philosophy of G. E. Moore* (533–677). New York, NY: Tudor Publishing.

Myin, E., and Hutto, D. D. (2015). 'REC: just radical enough'. *Studies in Logic, Grammar and Rhetoric* 41(54), 61–71.

Rietveld, E., and Kiverstein, J. (2014). 'A rich landscape of affordances'. *Ecological Psychology* 26, 325–352.

Ryle, G. (1949/2009). *The Concept of Mind*. New York, NY: Routledge.

Searle, J. R. (1995). *The Construction of Social Reality*. New York, NY: The Free Press.

Stapleton, M. (2013). 'Steps to a "properly embodied" cognitive science'. *Cognitive Systems Research* 22–23, 1–11.

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard University Press.

Thompson, E., and Stapleton, M. (2009). 'Making sense of sense-making: reflections on enactive and extended mind theories'. *Topoi* 28, 23–30.

Tuomela, R. (2007). *The Philosophy of Sociality: The Shared Point of View*. New York, NY: Oxford University Press.

Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: The MIT Press.

von Wright, G. H. (1963). *Norm and Action*. London, UK: Routledge & Kegan Paul.

Wedgwood, R. (1999). 'The a priori rules of rationality'. *Philosophy and Phenomenological Research* 59(1), 113–131.

Wedgwood, R. (2002). 'The aim of belief'. *Philosophical Perspectives* 16, 267–297.

Wilson, M. (2002). 'Six views of embodied cognition'. *Psychonomic Bulletin & Review* 9(4), 625–636.

Wilson, R. A., and Clark, A. (2009). 'How to situate cognition: letting nature take its course'. In P. Robbins and M. Aydede (Eds.), *The Cambridge Handbook of Situated Cognition* (55–77). New York, NY: Cambridge University Press.

Wright, C. (1980). *Wittgenstein on the Foundation of Mathematics*. London, UK: Duckworth.

Wittgenstein, L. (1953/1958). *Philosophical Investigations*. Trans. by G. E. M. Anscombe. Prentice Hall, NJ: Basil Blackwell.

Wittgenstein, L. (1958). *The Blue and Brown Books*. Oxford: Clarendon Press.

Zangwill, N. (2005). 'Normativity and the metaphysics of mind'. *Australasian Journal of Philosophy* 88(1), 21–39.

Zangwill, N. (2010). 'The normativity of the mental'. *Philosophical Explorations* 8(1), 1–19.

Ziemke, T. (2003). 'What's that thing called embodied cognition?' *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum. Online access: http://www.skidmore.edu/~flip/Site/Lab/Entries/2007/1/25_Embodiment_files/cogsci03.1.pdf.

Paper 1

Cosmos and History: The Journal of Natural and Social Philosophy, vol. 9, no. 1, 2013

SITUATING NORMS AND JOINTNESS OF SOCIAL INTERACTION[1]

Patrizio Lo Presti

ABSTRACT: The paper argues that contexts of interaction are structured in a way that coordinates part actions into normatively guided joint action without agents having common knowledge or mutual beliefs about intentions, beliefs, or commitments to part actions. The argument shows earlier analyses of joint action to be fundamentally flawed because they have not taken contextual influences on joint action properly into account. Specific completion of earlier analyses is proposed. It is concluded that attention to features distributed in context of interaction that signal expected part actions is sufficient for a set of part actions to qualify as a joint action.

KEYWORDS: Social Interaction; Joint Action; Status Functions; Norms

1. INTRODUCTION

People walk, dance, and sport together, as friends, family, colleagues, fellow citizens, and so on. Social interaction occurs everywhere–around dinner tables, in rush hour traffic, at work, and in train carriages. People coordinate actions, at red lights, in queues, and go to the ballot the same day. People are committed to action procedures in interaction, avoiding collisions, not pushing into queues, leaving space for hurrying ambulances, and so on. When we together perform actions by means of each performing a part we engage in *joint action*.

But under what conditions do sets of part actions qualify as joint action? I will review two camps of earlier analyses of joint action (sections 2 and 3). The important

[1] This research was funded by the ESF programme EuroUnderstaning as part of the NormCon group.

difference between them is what conditions they state about how agents´ mental states must be related for their part actions to qualify as a joint action. According to some it is necessary that we have mutual beliefs about what part actions we intend to perform, about what is our goal, or about our commitments to perform certain part actions. Others reject the notion that such mental state relations are necessary, at least for some kinds of joint action. For some kinds of joint action it has been claimed that it is sufficient for sets of part actions to qualify as joint if agents realise that they cannot perform the whole action themselves and that they attend to each other in a special way. I find both camps fundamentally flawed and in dire need of completion.

I propose that context coordinates part actions. Specifically I propose re-evaluation of analyses of joint action in light of the 'proposition of situatedness' that, *agents acting together are necessarily situated in social contexts structured around rules, roles, and functions, which are signalled by objects distributed in context of interaction*. A consequence of taking the situated perspective is that the class of joint actions is wider than if restricted to common knowers, mutual believers, and joint attenders.

What about context of interaction? Echoing Searle, we assign status functions, which entail power relations, to objects and people and accept objects to count as indicators of such functions and power relations. We distribute functions and roles in social environs and invest objects with signalling functions about which actions are expected, which are the norm in the sense of being permitted or prohibited. The proposition of situatedness finds support in Searle´s social ontology (section 4) and can be cashed out provided an understanding of how people attend to features of context of interaction that signal which actions are expected, the norm, and allowed or not allowed (section 5). That functions distributed in contexts of interaction influence joint action has repercussions for all earlier analyses of joint action.

2. MAXIMALISM

The first camp of analyses of joint action to be re-evaluated has been called 'maximalism' (Pacherie 2011). 'Maximalist analyses' refers to analyses that state as necessary for part actions to qualify as a joint action that they are performed either in common knowledge or mutual belief about intentions, goals, and reasons for action. The proposition of situatedness, properly developed, will show that part actions qualify as joint action without participants´ mental states being related in ways stated as necessary by maximalists. It will also show that maximalism is inaccurate explanation of commitments´ involvement in joint action. 'Maximalist explanations of commitments´ involvement in joint action' refers to any explanation tying commitments´ involvement in joint action to participants´ mutual belief or common knowledge about intentions, beliefs, or commitments to part actions.

First question, *How can set of actions qualify as joint action without mental states being related in ways stated as necessary according to maximalists?* To see how, we need know more about maximalist necessary conditions for actions to qualify as joint.

The first necessary condition that characterise maximalism is common knowledge. Part actions x_1 and x_2 qualify as a joint action X if and only if they are performed from common knowledge. Common knowledge is necessary for agents to engage in joint action according to Gilbert (1989, 1990, 2001, 2003, 2006, 2007a, 2007b), Bratman (1992, 1993, 2009), and Miller (2007). According to these authors, agents have common knowledge of some fact if each knows the fact and knows that each knows the fact, and so on. What fact must be commonly known is, according to Gilbert (1989, p. 189) that both have expressed readiness to engage in joint action, or, in a way we will come back to, that they are jointly committed to a shared intention (2000, p. 22). According to Bratman (1992, p. 335) it must be common knowledge that both have uncoerced intentions to engage in joint action.

Weaker variants of maximalism states as necessary for part actions to qualify as joint action that part actions are performed from mutual belief about intentions and beliefs relevant for effectuation of the whole action (Tuomela 1993, 2005, 2006, 2007; Tuomela and Miller 1988; Haakli, Miller, and Tuomela 2010; Pettit and Schweikard 2006; Pettit and List 2011). If we replace 'know' with 'believe' in the above definition of common knowledge, then we have the maximalist definition of mutual belief. Pettit and Schweikard (2006, p. 23) say that it is necessary for actions to qualify as joint action that each agent believe that others intend to do their parts and that agents mutually believe that this is the case. Tuomela (2007, p. 112) says that it is a necessary condition that agents act partly because it is mutually believed that they jointly intend to perform the joint action.

By distinguishing two contexts of interaction we can now begin to glimpse the force behind the proposition of situatedness and why maximalist analyses of joint action should be re-evaluated. What will distinguish the two contexts is that in the first there is no object signalling what part actions are expected, while in the latter there is such an object. Suppose first you are out driving. The driver behind you starts flashing his headlights. You wonder what you are supposed to do–what he or she intends, what he or she believes you have not noticed, etc. Perhaps your boot is open, or your blinkers on. This is a confusing situation, and a maximalist can explain why: you and the driver behind you will have problems acting jointly or to appropriately coordinate part actions because your mental states are not appropriately related. But now suppose you notice that the car behind you is an ambulance. So you veer to let it pass. This is not a confusing situation. Why is action procedure suddenly not

confusing? The only relevant distinguishing feature is that in the latter situation your attention to certain features of context, e.g., car lacquer and coloured light, indicate appropriate action procedure. In the second situation you are able to coordinate part actions not because you know or have beliefs the contents of which refer to the other´s mental states but because features of context signal what part actions are expected. This is as predicted by the proposition of situatedness. The example here is meant only to open for a point of entry of a situated perspective on joint action, a blind spot in maximalist analyses. Situatedness will be elaborated later on.

Second question, *Why is maximalism inaccurate explanation of commitments´ involvement in joint action?* Let´s review some maximalist explanations of commitments´ involvement in joint action.

Some maximalists say that agents will be committed to part actions in joint action. According to Gilbert (2009, p. 179; cf. 2000, p. 18) "Members of some population P *share an intention* to do A if and only if they are jointly committed to intend as a body to do A," where it is necessary for joint commitment 'as a body' that agents under conditions of common knowledge have expressed readiness to be individually committed to the joint action (2007b, p. 10; 2009, p. 180). When one is committed in this sense one is expected to carry out the intention accordingly and not to change one´s mind. Failure to do so gives others right to rebuke. Pettit and Schweikard (2006, p. 33) say that, "it is in virtue of their acting together that people are jointly committed to one another." Both explanations of commitments´ involvement in joint action are thus maximalist explanations–commitments are involved only if there is common knowledge about commitments to part actions or part actions are performed because agents mutually believe that others will perform their parts.

Consider the ambulance example again. Are you committed to let the car behind you pass when you notice that it is an ambulance? What happens in this situation is that you notice some feature, e.g., car lacquer, which signals what you are supposed to do. It is not necessary that you and the particular other in the context of interaction–the driver of the ambulance–suddenly form beliefs about each others´ intentions, beliefs, or commitments to part actions. Indeed, the ambulance might be on autopilot, hijacked, or controlled remotely by satellite–you will still know what action procedure is expected and part actions will be coordinated according to expectations, provided of course that each agent individually intends to act in accordance with expectations.

The crucial point of pivot in both ambulance-examples is this: *The only relevant difference between the scenarios with a civilian and an ambulance is that in the latter you attend to some feature of context that signals expected action procedure*. Appropriate coordination of part actions and commitments to action procedures in neither case necessitates reference

to others´ mental states in content of intentions or beliefs of either agent. The proposition of situatedness embraces what maximalist analyses of joint action and explanations of commitments eschew–that is, features of context of interaction that signal to situated agents what roles and functions of persons and objects apply will influence the interaction. Before elaborating further the situated perspective, let´s discuss 'minimalist' analyses of joint action.

3. MINIMALISM

What distinguishes maximalists and minimalists is that minimalists reject the notion that common knowledge or mutual belief is necessary for sets of part actions to qualify as joint action. Joint actions focused by minimalists are of 'simple kind'. Simple kind joint actions are, e.g., carrying a two-handled object together a shorter distance, early in development joint action, and emotionally guided joint action.

From the perspective of situatedness minimalists react to maximalism for reasons that miss something central about joint action–they do not consider how context influences joint action. Next, I provide a summary of minimalism.

For simple kind joint actions, Sebanz et. al. (2006, p. 70) suggest that shared goal representations and capacity to predict others´ part actions´ effects are sufficient conditions for sets of part actions to qualify as joint action. These conditions can be met if agents attend to each other, each others´ actions, and the ambient. Vesper et. al. (2010, pp. 999-1001) further specify building blocks for joint action. These building blocks are: first, that agents realise that own task alone is insufficient for the whole action; second, capacity to monitor and predict own and complementary tasks´ approximation to goal; and third, coordination smoothers such as exaggerated movements the goals of which are easy to predict. Findings in neuroscience show that agents, in turn-taking studies where other´s actions are perceptually available, exhibit increased action response prohibition during observation of other´s turn (Tsai et. al. 2005), suggesting that action observation triggers shared goal representations (cf. Knoblich and Jordan 2002, 2003).

Michael (2011) elaborates coordination smoothers to include expression (verbally or not) of emotion, which indicate what action one is prepared to perform. Anger, blushing, and sighing, for instance, indicate tractability of action procedures. Brownell (2011) draws upon findings in developmental psychology suggesting that children engage in joint action and understand commitments in interaction supposedly prior to development of cognitive abilities demanded by maximalist analyses (Tomasello et. al. 2005; Rakoczy 2006; Tomasello and Carpenter 2007; Tomasello and Rakoczy 2007; Rakoczy, Warneken, and Tomasello 2008; Carpenter 2009; Gräfenhain et. al. 2009).

Brownell argues that early joint action does not meet maximalist conditions. She reasons that infants are able to act jointly due to adults´ scaffolding and engaging actions to which the infant attends and has an interest in. Pacherie (2011) suggests that it is much easier to act as a 'we' than Gilbert claims. According to Pacherie it is sufficient that an agent 'frames' him or herself as a 'we' together with others and team-reasons in Bacharach´s (2006, p. 121) sense that "somebody 'team-reasons' if she *works out the best possible feasible combination of actions for all the members of her team, then does her part in it.*" Whether team reasoning is induced depends on circumstantial features that stimulate group identification, not on inferring whether others team-reason (Pacherie 2011, pp. 186, 189).

Kutz (2000, p. 11) emphasises that it is a sufficient condition for sets of part actions to qualify as joint action that agents have 'participatory intentions.' A participatory intention is an individual´s fostering of a collective end by means of own action and "requires neither positive beliefs about others´ intentions nor dispositions of responsiveness" (p. 20). Searle can be classified minimalist. This may raise some eyebrows since Searle demands that agents have 'collective intentions' to engage in joint action. But Searle (1990, p. 406; 1995, p. 24; 2010, p. 46; cf. 2002) denies that any amount of mutual beliefs or common knowledge in performance of part actions add up to joint action. I will have more to say about Searle´s account later.

An analysis of joint action is 'minimalist' in the present sense if it states that it is not necessary for part actions to qualify as joint that there is common knowledge or mutual belief about intentions, beliefs, or reasons for action. Preceding paragraphs summarize minimalism.

Positive minimalist accounts of join action differ. We have seen that some claim that action observation, which trigger shared goal representations, together with joint attention suffices for simple joint actions. Brinck and Gärdenfors (2003) claim that joint attention suffices for sets of actions to qualify as joint action if the goal does not have to be planned for. Joint attention is achieved when agents mutually attend to each others´ states of attention, make attention contact, and alternate gaze between each other and a shared object of attention (Brinck 2001, pp. 268-70, 2004, p. 196). Searle invokes the 'thesis of the Background,' that there "is a set of nonrepresentational mental capacities that enable all representing to take place" (1983, p. 143), and claims that a "Background sense of the other as a candidate in cooperative agency" (1990, p. 415) plus collective intentionality suffices for joint action.

Although maximalism and minimalism are very different with respect to how agents´ mental states must be related in performance of part actions for those actions to qualify as joint action I would like to focus on a common feature between them. Common to minimalists and maximalists is failure to ask in what way features of

context of interaction coordinate or at least influence coordination of part actions of agents situated in context. From a situated perspective we will see that, even for kinds of joint action focused by minimalists, lower levels of attention than joint attention suffices for joint action.

The review of maximalism and minimalism in this and the last section serves to introduce problems with current analyses of joint action. The problems stem from not taking contexts´ influence on joint action seriously.

These are no minor problems. I have used an example with an ambulance, but this is just the tip of an iceberg of situations problematic for maximalism and minimalism. It will become clear in the next two sections that counterexamples to earlier analyses teem when we take a situated perspective on joint action. Think of any context of interaction and ask yourself whether it contains objects that signal to agents situated in that context what actions they are expected to perform and what actions are norm. Think of wedding rings, police badges, or dress codes in different contexts–at church, on a date, or at a crime scene. Now ask yourself, *What happens with interaction if in any of the contexts you imagined you add or remove such an object?* The answer is that coordination of part actions is influenced when such objects are present and attended to in this sense: agents involved need not know or have beliefs the content of which refer to particular others´ mental states in the context of interaction.

Having reviewed earlier analyses of joint action and explanations of commitments´ involvement, the entry point of the situated perspective in the debate about when sets of part actions qualify as joint action is clear. The situated perspective departs from an 'outside-in' point of view, focusing how *contextual factors* can replace individuals´ internal representations of each others´ mental states in explanations of joint action. From the situated perspective the 'inside-out' point of view, focusing *mental state relations* as enabling conditions for joint action appears flawed and incomplete. To further spell this out I turn next to analyse how contexts of interaction can be structured around roles, rules, and social functions that, if embodied in objects in context, signal appropriate actions procedures in social interaction.

4. CONTEXTS OF INTERACTION

Searle´s (1995, 2010) social ontology is handy to account for how contexts can be structured to influence social interaction and coordinate part actions into joint action. Other social ontologies may serve to illuminate situatedness, but Searle´s is well known. Which ontology we choose is ultimately a matter of taste, as I will show later. Bicchieri´s (2006) analysis of norms in conjunction with Searle´s social ontology can explain commitments´ involvement in joint action from a situated perspective.

Social interaction is always performed in context–'context of interaction' for short. Most contexts of interaction are structured so that roles, rules, and social functions apply to persons and objects context. How is such structuring of contexts of interaction to be understood? First out, consider Searle´s notions of *imposition of status function.*

Status functions are functions of objects or individuals accruing not by virtue of physical constitution but by virtue of status function declared to apply. Status function declarations´ propositional content represents a state of affairs that does not yet exist, e.g., "owners of such-and-such pieces of plastic count as citizens", such that if others accept the declaration it is brought into existence (2010, pp. 85, 93-5). Status function declarations have the double direction of fit. If the piece of plastic endows its owner with the status function "counts as citizen" by virtue of the declaration (word-to-world direction), and if the owner acquires, through others´ acceptance, the status function it is represented as having, i.e., "citizen" (world-to-word direction) then the declaration is successful and the status function created.

Secondly, since status functions are not continuously imposed from context to context we want to know how they *persist*. Status functions can persist by being assigned by *standing* status function declarations. Standing status function declarations are declarations with an added proviso, e.g., "unless otherwise declared" or "from hereon" (Searle 1995, p. 97). Contracts and laws are paradigm standing status function declarations. For status functions to persist, whether declared as standing or not, it is necessary that they are *continually collectively recognised.* Collective recognition is an attitude of agents toward the propositional content of the original declaration. The form of collective recognition is, e.g., "We recognise that owners of such plastic cards count as citizens" (cf. p. 103). Continuation of collective recognition ensures persistence of status functions, why reiteration of imposition is not necessary for, e.g., maintained citizenship.

Thirdly, and importantly for explaining commitments in joint action, we need to understand the relationship between status functions and rights, duties, obligations, etc., in contexts in which they apply. Status functions entail deontology. Deontology is the set of entitlements, obligations, prohibitions, etc., accruing to objects or persons by virtue of their status functions. Deontology opens up a range of actions of people and uses of objects by virtue of the imposition of function. For instance, you are not only entitled to vote if you count as citizen, you are also obliged to pay taxes. Objects or people with a status function must (not) or may act or be used in certain ways. The reason why objects or persons must (not) or may act or be used in certain ways is that the status functions that accrue to them are collectively recognized, and thus the

range of actions allowed, prohibited, obligatory, etc., put others in position to expect certain action procedures in context of interaction (Searle 1995, p. 23).

Fourth, and importantly for understanding how status functions can be attended to, status functions are often indicated. There are indicators of status functions, and thus indicators of what roles, rules, and functions apply in context–'indicators of deontology' for short. Indicators are usually perceivable objects. Identity cards, uniforms, badges, and signatures are sample indicators (Searle 1995, p. 119f). Although one cannot see on an indicator of deontology that a person is allowed or prohibited from acting in some way, it signals such deontology by being collectively recognised to so signal–if there were no collective recognition there would be no deontology to signal to begin with. Importantly here, there are objects, collectively recognised as indicators of status functions, signalling in context of interaction what roles, rules, and functions apply. By implication, there are objects indicating what range of actions is open as allowed, prohibited, obliged, and so on, by agents with certain status functions in context of interaction. (There is, of course, only *one* function of indication, but it indicates both the presence of status function and the range of actions the status function opens up).

Together imposition and persistence of status functions, the deontology they entail, and indicators of them explain what is meant by 'structure of context of interaction'. With analysis of context of interaction at hand we can now provide a situated perspective on commitments´ involvement in joint action.

Commitment in joint action will here mean that one has obligations to carry out part actions and that others have reason to expect that one carries them out–that is 'commitment' in Gilbert´s sense–or that one is responsive to help others perform their part actions if necessary for the joint action to succeed, in Bratman´s (1992) sense. Commitments to joint action, then, involve expectations that others will carry out their actions on pain of sanctions, or that one will be helped if one has problems carrying out one´s own part, or both.

There is a definition of norms that is reminiscent of these notions of commitment. Social norms, defined as some rule´s application in context, preference for conformity, belief that the rule is collectively recognised, and belief that conformity is expected and that deviance may evoke negative reciprocation (Bicchieri 2006, p. 11) concur with deontology in context of interaction as defined above. That is, status functions entail deontology–permissions, obligations, prohibitions, etc., that persist by continuation of collective recognition. Since deontology in contexts of interaction is by necessity collectively recognised, expectations about action procedure will be shared in each collective with the same deontology. People sharing deontology who

interact in a context where deontology is present and indicated will be able to coordinate part actions normatively appropriately by attending to features of context indicating roles and functions relevant to their interaction. Bicchieri (2006, p. 73) puts this nicely:

> People are able to coordinate actions and expectations despite limited access to the operation and content of their and others´ minds … Such coordination is possible because people share collective perspectives which have led them to develop similar inferences and interpretations of common situations, object, and events.

Crucial to differentiate the situated explanation of commitments in joint action from maximalist explanations, it is according to the former not necessary that agents´ mental states are related in ways stated as necessary by maximalists for commitments to be involved in joint action. Commitments are involved because agents have expectations about action procedure. These expectations´ contents need not refer to others´ mental states. How does this work? To answer this, the next section explicates what I mean by attention to indicators of status functions and shared expectations about action procedure.

5. ATTENTION IN ACTION

If agents attend to indicators of status functions in context of interaction in which they are situated joint action will be influenced because the indicators signal what part actions are expected and the norm. What is meant by 'attention to indicators of status functions'? Well, what is meant by 'attention'?

Let attention be modality neutral. Objects and events assigned social functions, roles, etc., in contexts of interaction, if attended to, signal appropriate action procedures to situated agents. For instance, when you enter a friend´s house you search for a coat hanger, when in the supermarket people in uniform can give information, when you are out driving the sound of sirens alerts you to take appropriate action. Importantly, when you attend to features of context action possibilities become salient. If you attend to features of which you have prior experience about what collectively expected actions they signal, then you will recognise action possibilities in present context by categorisation of present input from attention (Brinck 1997, 1999, 2001, 2004; Brinck and Gärdenfors 1999; Brinck, Zlatev, and Andrén 2006; Bicchieri 2006).

Status function, we have seen, have indicators–e.g., wedding rings, badges, brooches, car lacquer, the chair man´s club, uniforms, etc. Status functions, furthermore, are by necessity collectively recognised and have concomitant

collectively recognised deontology. *What happens, then, if we attend to indicators of status functions?* The action procedures signalled will be such that are collectively recognised and expected. Indicators of status functions trigger expectations about and commitment to collectively recognised action procedures when attended to.

We can now say this: indicators of status functions in context of interaction, if attended to by situated agents, trigger expectation about collectively recognised and expected action procedures. But, this is still not an account of how contexts of interaction influence and facilitate coordination of *joint* action. To facilitate joint action the expectations indicators trigger must be of a special kind: they must be shared expectations. What is meant by 'shared' expectations?

Expectations are shared if their contents are the same for two or more agents. For instance, expectations with the content "*A* will stop by the side of the road" are shared if two or more agents expect that *A* will stop by the side of the road. But, merely sharing expectations is not sufficient for joint action to be guided. People can share expectations that it will rain at the other side of the globe tomorrow without this facilitating joint action. Let´s therefore add, what is core of the situated approach, that the content of shared expectations must refer to an action procedure that is collectively recognised as expected in the context of interaction in which the agents are situated. However, *that action procedure A is collectively recognised as expected* does not have to figure in the content of shared expectations. It suffices that *action procedure A* is triggered by attention to an indicator of deontology and that agents situated in the context act accordingly. From these elaborations of 'attention to indicators of deontology' and 'shared expectations' we reach the backbone of the situated approach to joint action.

Agents attending to indicators of status functions, which entail deontology, in contexts in which they are situated, will share expectations about action procedure, because deontology *is* collectively recognised obligations, prohibitions, permissions, etc., about action procedure. Agents will be committed to certain actions in context of interaction, because they share expectations about allowed or not allowed action procedures. Importantly, it is not necessary for expectations about action procedures to be shared that their contents refer to other agents´ or the collective´s expectations or mental states. Status functions do not coordinate agents´ beliefs or knowledge about each others´ mental states. Status functions in context of interaction, through indicators of deontology to which people attend in action, coordinate part actions by means of which joint actions are performed. However, part actions performed will not qualify as joint by mere coincidence. Part actions will be performed as parts of collectively expected action procedures in the context of interaction, given of course

that agents intend to perform what their attention to features of context indicate as appropriate (if they do not, they breach commitments). It is therefore not necessary that agents interacting have beliefs the content of which refer to any particular other´s intentions, beliefs, or commitments to part actions in the context of interaction for part actions to qualify as joint action. So we can see that contextual factors can replace individuals´ internal representations of each others´ mental states in explanations of joint action.

This section and the preceding establish the proposition of situatedness*: agents acting together are necessarily situated in social contexts structured around rules, roles, and functions, which are signalled by objects distributed in context of interaction*. We are now in a position also to understand why, and justify the claim that, the class of joint actions is wider than as demarcated by maximalists and minimalists.

To abate any doubts about this, let´s return to the counterexamples that presented problems for maximalist analyses. The present account dissolves the problems highlighted in the counterexamples in a manner obscured also in the minimalist 'inside-out' perspective. We will see that the situated approach with its stress on context of interaction generates a flexible approach to analysing joint action suitable to complement earlier analyses.

First, then, consider the case with the ambulance, where in one scenario there is no indicator of status functions while in the latter there is. The driver behind you starts flashing his headlights. You wonder what you are supposed to do–what he or she intends, what he or she believes you have not noticed, etc. One way to explain aggravation of appropriate performance of part actions here is focusing disconnectedness of mental state-relations. Specifically, the focus would be on lack of reference to other´s mental states in the content of beliefs or intentions from which respective actions are performed. This is the maximalist-minimalist explanation.

Now suppose you notice that the car behind you is an ambulance. So you veer to let it pass. *The only difference is that you suddenly attend to something that signal appropriate part actions*. In the second situation part actions are coordinated not because you know or have beliefs the content of which refer to the other´s mental states but because features of context signal what part actions are expected. Of course your mental states *may* be related in the way necessary according to maximalists, but it is *not* necessary, and of course your attentional states *may* be connected in the minimalist sense. But, noticing that circumstances for interaction turn on contextual factors–and I dare say that failure to notice as much robs us of fruitful perspective on joint action–motivates shifting attention from shared mental content to shared contexts of interaction. In the situated perspective it suffices that you attend to indicators that signal and trigger

expectations about appropriate action procedure–that is, expected and recognised as allowed, prohibited, obligatory, etc.

To repeat, it is no mere coincidence what part actions are performed. Agents´ mental states in a sense *are* related on the present account. But they are not related in the maximalist sense of reciprocal referencing. Mental states–expectations about action procedure–need stand in correlation to a deontology. They are shared in virtue of being so correlated.

The situated account of joint action thus explains how part actions are coordinated in joint action without maximalist necessary conditions for mental state relations being met. But is situatedness different from minimalism? It is.

Minimalist analyses of joint action, we have seen, state that, for joint actions of simple or early in development kind, or for joint action guided by expressions of emotions, maximalist conditions for joint action are not necessary. From a situated perspective, in contrast, also for non-simple joint action attention to indicators of status functions is sufficient for part action coordination into joint action. For simple kind joint action we saw that some minimalists claim that action observation, shared goal representations, and joint attention is sufficient for part actions to qualify as joint action (Vesper et. al. 2010; Sebanz et. al. 2006). From a situated perspective, in contrast, even for simple kinds of joint action attention to indicators of status functions is sufficient for sets of actions to qualify as joint action. In this respect the situated account requires only lower levels of attention (Brinck, Zlatev, and Andrén 2006) for part actions to be coordinated into joint action. And in this respect the situated account concerns kinds of joint action focused by both minimalists and maximalists. Brinck and Gärdenfors do say that whether joint attention is sufficient for part actions to be appropriately coordinated is altogether a question of context (1999 p. 94, 2003 p. 489; Brinck 1997 p. 130, 2001 p. 263). I agree with that statement. However, situatedness accounts for how objects in context signal appropriate part actions, something Brinck and Gärdenfors do not account for.

Maximalism faced a second problem. The problem was to explain commitments' involvement in joint action without detour to believed sharing of mental content.

Are you committed to leave space once you notice the car behind you is an ambulance? Yes, because there is a collectively recognised deontology, e.g., ambulances transport the acutely sick and must be allowed priority in traffic. The deontology was indicated, e.g., by the car lacquer or sounding of sirens. Attention to these indicators triggered expectations about allowed, prohibited, and obligatory action procedure. The two of you need not have any beliefs about each others´ beliefs, intentions, or commitments to part actions. It is sufficient that indicators trigger

expectations about action procedure. Commitments are involved because features of context signal appropriate action procedures. For this to be the case it is not necessary that participants internally represent the contents of each others´ mental states.

In conclusion, first, features of context to which agents attend in action can replace reciprocal mental state referencing in part action performance for part actions to qualify as joint action. Second, part actions so coordinated are not incidental but qualify as joint because they are performed in line with a collectively recognised deontology. Third, commitments can be involved in joint action without meeting maximalist conditions. Fourth, an analysis of the structure of contexts of interaction and attention in action by agents situated in contexts allow explanation of part action coordination and part actions´ adding up to joint action.

I hope to have shown why we need to take context and the proposition of situatedness seriously in analysing joint action, what problems it solves to do so, and how we can do it. It is important to remember that the situated perspective is not meant to *replace* either maximalism or minimalism. It is meant to be *complementary*. Indeed, situatedness *cannot*, as here formulated, provide an independent analysis of joint action or explanation of commitments´ involvement in joint action. Why this is so is explained in the next section, where several objections to situatedness are considered.

6. DEFENDING SITUATEDNESS

I will consider seven objections threatening situatedness. The first concerns whether it accounts for joint action at all, the second that the account is circular. The third to fifth objection are against the analysis of contexts of interaction, and the sixth against the reading of earlier analyses of joint action. The last objection is that the proposition of situatedness is trivial.

(1) First objection is that part actions that are not properly joint will qualify as joint action on the situated account. For instance, simultaneous, spatiotemporally proximal, or mere incidental part actions appear to qualify. Are the conditions stated as sufficient for part actions to qualify as joint by situationalists too admissive? The answer is: no. From attention to indicators of status functions in context of interaction specific individual intentional actions are performed which correspond to the range of actions open as allowed, prohibited, obligatory by deontology that applies in context. Actions are coordinated by context but also by recognition of expectations in context of interaction by agents situated in context. So, sets of part actions counted as joint actions according to the situated approach are no mere simultaneous, spatio-

proximal, or incidental actions. Situated agents act as agents related in a special way by context of interaction and expectations about action procedures without the contents of those expectations necessarily referring to each others´ mental states. The conditions stated as sufficient for sets of part actions to qualify as joint according to situationalists are thus not too admissible. The account distinguishes mere synchronous, spatiotemporally proximal, similar, and chance coordinated actions from part actions coordinated into joint action.

(2) Second objection is that for context of interaction to coordinate part actions into joint action in the situationalist sense contexts must be appropriately structured around status functions, but status function creation and persistence requires continual collective recognition of the status functions. On pain of circularity, at some point it cannot have been a sufficient condition for part actions specifically intended to *jointly impose* status functions that agents attended to indicators of status functions that coordinated part actions and triggered shared expectation about action procedure. Can the situationalist avoid this objection? In the previous section I pointed out that situationalism is no *replacement* but rather a *complement* of earlier analyses. What conditions we state as necessary or sufficient for sets of part actions to qualify as joint action depends on the kind of action under consideration. The situationalist has to admit that there are some kinds of joints action that situationalism cannot explain. There are at least two such kinds. First, at some point in time conditions stated as sufficient on the situational account for part actions to qualify as joint action were not sufficient for joint *creation* of status functions. However, it is still true that contexts of interaction suitably structured and attended to by situated agents can coordinate part action into joint action of kinds focused by both minimalists and maximalists (section 5). Furthermore, for status functions to *persist*, once created, it is not necessary that agents recognise that others recognise that they recognise, and so on, that there is collective recognition of status functions. Secondly, situationalists cannot account for instances of joint action early in development when children supposedly do not recognise that, or which, deontology is indicated by certain objects. But, the conclusion that both maximalism and minimalism need completion from situationalism, or some similar account, still follows.

(3) Another objection might be that Searle´s social ontology has problems on its own (Johansson 2003; Meijers 2003; Zaibert 2003; Smith 2003) and cannot support situationalism. As an objection to situationalism this is beside the point. Ultimately, an alternative social ontology will have to account for the existence and persistence of facts that influence social interaction, such as roles and functions of agents and objects situated in contexts of interaction. An alternative social ontology will have to explain

also the emergence of rules and norms of interaction and how agents situated in contexts of interaction are able to identify which rules, roles, functions, and norms apply. As far as these social phenomena are not explained *away* by an alternative social ontology, compatibility with the general situationalist credo is preserved. It just so happens, as a contingent empirical fact at the very least, that people have different roles, which entail ranges of expected actions in different contexts of interaction. Searle´s social ontology is at the moment an influential genealogy of the structure of our social world.

(4) The last answer leads us to a fourth objection. Some might say that contexts of interaction structured in the Serlean sense is not what coordinate part actions, even if we admit that features of contexts can facilitate coordination of part actions into joint action without involvement of believed sharing of mental states. We can explain, the objection might go, contextual facilitation of joint action by appeal to other social phenomena, e.g., conventions. Situationalism is compatible with the view that conventions facilitate joint action. Even a social ontology based *wholly* on conventions motivates a situationalist perspective on joint action according to which recognition of presence of convention in context of interaction coordinate part actions into joint action. Conventions, though, rely on common knowledge (Lewis 1969, p. 78), which would make this brand of situationalism maximalist. But conventions as well as status functions are indicated (p. 61) and salience and precedents can trigger sharing of expectations and facilitate coordination of part actions (p. 57). Therefore conventionally structured contexts of interaction can facilitate joint action as well as contexts of interaction structured around status functions.

Objections (3) and (4) point to a central virtue of situationalism. Situationalism is theory neutral. That is, whatever analysis or genealogy we adopt to explain the structure of social reality, the situationalist outside-in perspective on social interaction will have a say in analysing joint action, as complementation to the inside-out perspective of maximalist-minimalist views. More about this under objection (7) about triviality.

(5) Another objection to the adoption of Searle´s social ontology is that status functions are not always indicated (De Soto 2003). Furthermore, sometimes contexts of interaction are so complex in structure, or agents´ attentional access so limited, that joint action is complicated rather than facilitated. Some of the contexts in which we interact are hierarchically opaque, have immense circumference, or diffuse borders. Some status functions are barely indicated. But this point rather reinforces than erodes support for one of the consequences the situationalist proposition, the consequence that when we state conditions for part actions to qualify as joint action

we have to take contextual influence seriously. That contexts of interaction aggravate or facilitate joint action adds potency to the proposition that contexts of interaction influence coordination of part actions.

(6) There may be objections to my reading of earlier analyses (sections 2 and 3). For instance, Tuomela (2007, p. 191f) says that, because of presence of institutions, actions tend to become routine, and Searle (1995, 2010) might be thought to have implicitly noticed what has been argued in this paper. If it were true that Tuomela and Searle are already aware of the present argument that would be no objection. It would show that there is prior support for it. But it appears that neither Tuomela nor Searle have fully appreciated the consequences of the present account.

Tuomela maintains that action within institutions–that is, complexes of status functions–depends on agents´ acting as group members (2007, p. 198f). Acting as a group member requires that there is a social group, which is the case if and only if some individuals have accepted an ethos (goals, beliefs, standards, etc.) as constitutive for them as a collective, mutually believe that they are committed to the ethos, and mutually believe that they share beliefs that they are group members (p. 22). To act as a group member requires that it is mutually believed that there is an understanding that the ethos provides authoritative reasons for actions and that every member is committed to the ethos (p. 19f).

Since Tuomela requires that people´s mental states in institutionalised joint action still are appropriately related, he does not realise that features of context of interaction can coordinate part actions in a sense sufficient for joint action. On one occasion (2003, p. 156) Tuomela says, in parenthesis, that agents must act as group members "at least in the creation" of institutions. Now that statement would be in line with the above concession (objection (2)) that although joint *creation* of status functions necessitates stronger relations between mental states than sufficient for joint action on the situationalist account, joint action in contexts structured around *existing* status functions does not. However, Tuomela does not consider, what has been argued at length here, that features of contexts of interaction attended to by situated agents can replace mutual referencing of mental state contents in explaining part action coordination into joint action. Furthermore, situationalism is not restricted to joint actions within complex institutions. Think of any context of interaction *without* indicators of status functions. People might have to agree or have mutual beliefs about part actions and commitments to engage in joint action in such contexts. Now *add* to the context indicators that signal expected action procedure. What happens? *Part actions of agents situated in context become coordinated in correlation to deontology signalled by status function indicators distributed in context.* This is not appreciated in current analyses of joint action, minimalist or maximalist.

Regarding Searle´s theories of collective action (1990) and social ontology (1995, 2003, 2006, 2010), he does not seem to have appreciated how the latter can explain the former. Searle states, on the one hand, that collective action depends on individuals´ we-intentions to perform a collective action by-means-of or by-way-of individual action. On the other hand, status functions are created by status function declarations that are collectively accepted or recognised. Situationalism can be understood as an argument that these two parts of Searle´s legacy fit together in a way that he himself perhaps has not realised. The present argument consolidates Searle´s theory of social ontology and collective action by explaining how contexts, which are structured around status functions, influence joint action. The notion of attention that the situationalist uses is hardly an addition at all.

Situationalism not only consolidates but also seems to develop parts of Searle´s philosophical contribution since *Intentionality* (1983). Searle has been criticised for not explaining how agents engage in joint action without their mental states being related in the maximalist sense. He has been criticised for individualising collectivity by saying that a brain in a vat can we-intend a joint action. He has been criticised for appealing too soon to his thesis of the Background to explain how agents´ mental states are related in joint action. Searle´s appeal to a Background or embedded sense of the other as a candidate for cooperation (1990, p. 415; cf. 1983, p. 154) is his way, I take it, of saying that it is not mutual beliefs or common knowledge that make part actions joint but that part actions are performed from a mental state with certain psychological mode–a 'we-intentional' state. More follows from the thesis of the Background if coupled with the later social ontology. We invest our world with functions, Searle says. Those functions, he also says, persist because we continually collectively recognise them as existing, and they open up for ranges of actions. These actions are prescribed and proscribed and are collectively recognised as prescribed or proscribed. Objects signalling such functions, as is part of Searle´s social ontology, facilitate, the situationalist continues, coordination of part actions into expected and normatively appropriate joint actions. Such signalling objects if attended to therefore function as coordinators of part actions. Status functions define roles and functions and entail prescription and proscriptions, which are signalled by indicators. Indicators are collectively recognised to so signal and thus have a coordination function. So it is not only an embedded Background sense of the other that allows people to perform joint actions without mutual beliefs or common knowledge. What is more, in a slogan, we distribute the 'we' in our Serlean heads into contexts of interaction in which we are situated when we act together. By doing so we can act together by gathering from

the objects around us what we expect rather than from particular others´ mental states.

(7) Last objection before concluding, *Is the situationalist account of joint action trivial?* Everyone agrees that joint action is influenced by contextual factors but just find it too obvious to mention, right? First answer, no doubt contextual influence on joint action has been thought about before. Situationalism can be seen as clarification and specification of intuitions about contextual influences on social interaction and joint action. The account expresses such intuitions in a way that allows evaluation of them as theoretical complements to earlier analyses. It does so in a way that motivates re-evaluation of earlier analyses, and provides arguments for the intuitions. Second answer to the triviality-objection: situationalists do not rest content with the trivial credo that context is important. Situationalists aim at specific explanation of how, when, and why contexts influence social interaction and joint action. One consequence of situationalism is that we can adjudicate whether people were and predict if they will be able to act jointly. We can do so by examination of their situatedness. Were or will status functions be present? Were or will they be indicated? Were or will agents be able to attend to the indicators? Were or will the agents be members of the same or similar collectives with the same or similar deontology indicated by features of context? When we ask those questions we arrive at a fairly accurate answer to whether people were or will be able to engage in joint action in particular situations. This is not trivial. Situationalism provides credible predictions and explanations complementing earlier analyses. Taking a situated perspective on joint action is thus to appreciate that the structure of contexts of interaction *enable* participants' understanding of how to coordinate actions into normatively appropriate joint action. Third answer: it is baffling if maximalists or minimalists accept the proposition of situatedness but find it trivial and uninformative. Since context in the sense proposed render conditions stated by earlier analyses obsolete in a wide range of ordinary, everyday contexts of interaction, how can proponents of those earlier analyses accept such contextual influence and reduce it to triviality? Of course if, from a pre-theoretical stance, one were presented with maximalism, minimalism, and situationalism, one might find the latter trivial *if interpreted as stating merely that "contexts influence action"*. In that case it should be kept in mind that situationalism does not rest content with the sweeping claim "context matters", it articulates an approach with predictive and explanatory power that can be scrutinised, evaluated, argued for, and compared to other approaches beyond mere intuition. Furthermore, understanding how to coordinate is on this account situated in the stronger sense that contexts of interaction in which agents are situated enables normatively appropriate

coordination. Were we pre-theoretically to find analyses of joint action focusing mental content more appealing, then how should we explain the influence of external, material factors on joint action? The proposed situated perspective answers these questions.

7. CONCLUSION

The argument of this paper is simple: First, if contexts of interaction structured around status functions are common and influence joint action, then analyses of joint action must take contextual influence on join action seriously; contexts of interaction structured around status functions are common and influence joint action; therefore, analyses of joint action must take contextual influence seriously. Second, if an analysis of joint action does not take context seriously, then that analysis is at best incomplete. Third, contemporary analyses of joint action do not take context seriously. Therefore, these analyses of joint action are at best incomplete and in need of completion. Situationalism, suitably developed, provides that completion.

We can conclude that taking contextual influence on joint action seriously allows flexible account of joint action and commitments in joint action. Agents are situated in contexts invested with structure that signals appropriate coordination of and commitment to part actions. This is perhaps new, but part and parcel of Searle´s legacy. People walk, sport, and cook together, as friends, colleagues, and fellow citizens. They do so at home, at work, in rush hour traffic, at church, etc., sometimes with authority or in power relations. Whether the actions they perform are informed or motivated by believed sharing of beliefs, intentions, and commitments is one important factor in explaining their joint projects. Another important factor is the influence that the social environs in which they interact have on their expectations and identification of normatively appropriate action procedures.

Acknowledgements: I would like to thank Ingar Brinck, Johannes Brandl, Björn Petersson, Andras Szigeti, Johannes Roessler, Dan Zahavi, Ingvar Johansson, and Martin Viktorelius for useful comments on drafts and presentations of this paper.

Department of Philosophy and Cognitive Science
Lund University
Sweden
Patrizio.Lo_Presti@fil.lu.se

BIBLIOGRAPHY

Bacharach, Michael. *Beyond individual choice: teams and frames in game theory*. Princeton: Princeton University Press, 2006.

Bicchieri, Christina. *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York: Cambridge University Press, 2006.

Bratman, Michael. 'Shared Cooperative Activity', *The Philosophical Review* vol. 101, no. 2 (1992), 327-341.

Bratman, Michael. 'Shared Intention', *Ethics* vol. 104 (1993), 97-113.

Bratman, Michael. *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge: Cambridge University Press, 1999.

Bratman, Michael. 'Modest sociality and the distinctiveness of intention', *Philosophical Studies* vol. 144 (2009), pp. 149-165,

Brinck, Ingar. *The Indexical 'I': The First Person in Thought and Language*. Dodrecht: Kluwer Academic Publishers, 1997.

Brinck, Ingar. 'Attention and tool use in the evolution of language' in *Spinning Ideas: Electronic Essays Dedicated to Peter Gärdenfors on his Fiftieth Birthday* (1999), http://www.lucs.lu.se/spinning/.

Brinck, Ingar. 'Attention and the evolution of communication', *Pragmatics & Cognition* vol. 9, no. 2, (2001), 259-277.

Brinck, Ingar. 'Joint Attention, Triangulation and Radical Interpretation: A Problem and its Solution', *Dialectica* vol. 58, no. 2 (2004), 179-206.

Brinck, Ingar. 'Intersubjectivity and Intentional Communication', in Zlatev, J., Racine, T., Sinha, C. and E. Itkonen (eds.), *The Shared Mind: Perspectives on Intersubjectivity* (2008), 115-140. Amsterdam: John Benjamins Publishing Co.

Brinck Ingar. and Gärdenfors, Peter. 'Representation and Self-Awareness in Intentional Agents', *Synthese* vol. 118 (1999), 89-104.

Brinck, Ingar. and Gärdenfors, Peter. 'Co-operation and Communication in Apes and Humans'. *Mind & Language* vol. 18, no. 5 (2003), 481-501.

Brinck Ingar., Zlatev, Jordan and Andrén, Mats. 'An applied analysis of attentional intersubjectivity', *SEDSU Deliverable 7*, 2006.

Brownell, Celia. 'Early Developments in Joint Action', *Review of Philosophy and Psychology* vol. 2 (2011), 193-211.

Carpenter, Malinda. 'Just how Joint is Joint Action in Infancy?' *Topics in Cognitive Science* vol. 1 (2009), 380-392.

De Soto, Hernando. *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*. New York: Basic Books, 2003.

Gilbert, Margaret. *On Social Facts*. Princeton: Princeton University Press, 1989.

Gilbert, Margaret. 'Walking Together: A Paradigmatic Social Phenomenon', *Midwest Studies in Philosophy* vol. XV (1990), 1-15

Gilbert, Margaret. 'Collective Preferences, Obligations, and Rational Choice', *Economics and Philosophy* vol. 17 (2001), 109-119.

Gilbert, Margaret. *Sociality and Responsibility: new essays in plural subject theory*. Lanham, Maryland: Rowman & Littlefield Publishers, 2000.

Gilbert, Margaret. 'The Structure of the Social Atom: Joint Commitment as the Foundation of Human Social Behavior', in Schmitt, F. F. (ed.), *Socializing Metaphysics: The Nature of Social Reality* (2003), 39-64. Lanham, Maryland: Rowman and Littlefield Publishers, Inc.

Gilbert, Margaret. 'Rationality in Collective Action', *Philosophy of the Social Sciences* vol. 36, no. 1 (2006), 3-17.

Gilbert, Margaret. 'Searle and Collective Intentions', in Tsohatzidis, S. L. (ed.), *Intentional Acts and Institutional Facts: Essays on John Searle's Social Ontology* (2007), 31-48 Dodrecht: Springer.

Gilbert, Margaret. 'Mutual Recognition, Common Knowledge, and Joint Attention', in T. Rønnow-Rasmussen, B. Petersson, J. Josefsson, and Egonsson, D. (eds.), *Homage á Wlodek: Philosophical Papers Dedicated to Wlodek Rabinowicz* (2007), www.fil.lu.se/hommageawlodek.

Gilbert, Margaret. 'Shared intention and personal intentions', *Philosophical Studies* vol. 144 (2009), 167-187.

Gräfenhain, Maria, Behne, Tanya, Carpenter, Malinda and Tomasello, Michael. 'Young Children's Understanding of Joint Commitments', *Developmental Psychology* vol. 45, no. 5 (2009), 1430-1443.

Hakli, Raul, Miller, Kaarlo and Tuomela, RTaimo. 'Two Kinds of We-Reasoning', *Economics and Philosophy* vol. 26 (2010), 291-320.

Johansson, Ingvar. 'Searle's Monadological Construction of Social Reality', *American Journal of Economics and Sociology* vol. 62, no. 1 (2003), 233-255.

Knoblich, Günther. and Jordan, Jerome. 'The mirror system and joint action', *Advances in Cognitive Research* vol. 42 (2002), 115-134.

Knoblich, Günther and Jordan, Jerome. 'Action coordination in groups and individuals: learning anticipatory control', *Journal of Experimental Psychology* vol. 29, no. 5 (2003), 1006-1016.

Kutz, Christopher. 'Acting Together', *Philosophy and Phenomenological Research* vol. LXI, no. 1 (2000), 1-31.

Lewis, David. *Convention*. Oxford: Blackwell Publishers, 1969.

List, Christian and Pettit, Philip. *Group Agency: the possibility, design, and status of corporate agents*. New York: Oxford University Press, 2011.

Meijers, Anthonie, W. 'Can Collective Intentionality be Individualized?' *American Journal of Economics and Sociology* vol. 62, no. 1 (2003), 167-183.

Michael, John. 'Shared Emotions and Joint Action', *Review of Philosophy and Psychology* vol. 2 (2011), 355-373.

Miller, Suemas. 'Joint Action: The Individual Strikes Back', in Tsohatzidis, S. L. (ed.), *Intentional Acts and Institutional Facts: Essays on John Searle´s Social Ontology*, (2007), 73-92. Dodrecht: Springer.

Pacherie, Elisabeth. 'Framing Joint Action', *Review of Philosophy and Psychology* vol. 2 (2011), 173-192.

Pettit, Philip and Schweikard, David. 'Joint Actions and Group Agents', *Philosophy of the Social Sciences* vol. 36, no. 1 (2006), 18-39.

Rakoczy, Hannes. 'Pretend play and the development of collective intentionality', *Cognitive Systems Research* vol. 7 (2006), 113-127.

Rakoczy, Hannes and Tomasello, Michael. 'The Ontogeny of Social Ontology: Steps to Shared Intentionality and Status Functions', in Tsohatzidis, S. L. (ed.), *Intentional Acts and Institutional Facts: Essays on John Searle´s Social Ontology* (2007), 113-138. Dodrecht: Springer.

Rakoczy, Hannes, Warneken, Felix and Tomasello, Michael. 'The sources of normativity: Young children´s awareness of the normative structure of games', *Developmental Psychology* vol. 44, no. 3 (2008), 875-881.

Searle, John. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press, 1983.

Searle, John. 'Collective intentions and actions', in P. R. Cohen, J. Morgan and M. E. Pollack (eds.), *Intentions in Communication* (1990), 401-416 Cambridge: MIT Press.

Searle, John. *The Construction of Social Reality*. New York: The Free Press, 1995.

Searle, John. *Consciousness and Language*. Cambridge: Cambridge University Press, 2002.

Searle, John. 'Social Ontology and Political Power', in F. F. Schmitt (ed.), *Socializing Metaphysics: The Nature of Social Reality* (2003), 195-210. Maryland: Rowman and Littlefield Publishers, Inc.

Searle, John. 'Social ontology: Some basic principles', *Anthropological Theory* vol. 6, no. 1 (2006), 12-29.

Searle, John. *Making the Social World: The Structure of Human Civilization*. New York: Oxford University Press, 2010.

Smith, Barry and Searle, John. 'The Construction of Social Reality: An Exchange', *American Journal of Economics and Sociology* vol. 62, no. 1 (2003), 285-309.

Sebanz, Natalie, Bekkering, Harold and Knoblich, Günther. 'Joint action: bodies and minds moving together', *TRENDS in Cognitive Science* vol. 10, no. 2 (2006), 70-76.

Tomasello, Michael and Carpenter, Malinda. 'Shared intentionality', *Developmental Science* vol. 10, no. 1 (2007), 121-125.

Tomasello, Michael, Carpenter, Malinda, Call, Josep, Behne, Tanya and Moll, Henrike. 'Understanding and sharing intentions: The origins of cultural cognition', *Behavioral and Brain Sciences* vol. 28 (2005), 675-735.

Tsai, Chia-Cin, Kuo, Wen-Jui, Hung Daisy L. and Tzeng, Ovid J-L. 'An electrophysiology study of representing others´ actions', *Journal of Cognitive Neuroscience* Supplement: 70 (2005).

Tuomela, Raimo. 'What is Cooperation?' *Erkenntnis* vol. 38 (1993), 87-101.

Tuomela, Raimo. 'Collective Acceptance, Social Institutions, and Social Reality', *American Journal of Economics and Sociology* vol. 62, no. 1 (2003), 123-165.

Tuomela, Raimo. 'The We-Mode and the I-Mode', in F. F. Schmitt (ed.), *Socializing Metaphysics: The Nature of Social Reality* (2003), 93-128. Maryland: Rowman and Littlefield Publishers, Inc.

Tuomela, Raimo. 'We-Intentions Revisited', *Philosophical Studies* vol. 125 (2005), 327-369.

Tuomela, Raimo. 'Joint Intention, We-Mode and I-mode', *Midwest Studies in Philosophy* vol. XXX (2006), 35-58.

Tuomela, Raimo. *The Philosophy of Sociality: The Shared Point of View*. New York: Oxford University Press, 2007.

Tuomela, Raimo and Miller, Kaarlo. 'We-Intentions', *Philosophical Studies* vol. 53 (1988), 367-389.

Vesper, Cordula, Butterfill, Stephen, Knoblich, Günther and Sebanz, Natalie. 'A minimal architecture for joint action', *Neural Networks* vol. 23 (2010), 998-1003.

Woodward, Amanda, L. 'Infants´ ability to distinguish between purposeful behaviors', *Infant Behavior & Development* vol. 22, no. 2 (1999), 145-160.

Zaibert, Leo A. 'Collective Intentions and Collective Intentionality', *American Journal of Economics and Sociology* vol. 62, no. 1 (2003), 209-232.

Paper 2

abstracta Linguagem, Mente & Ação ISSN 1807-9792 Volume 7, Number 1, pp. 5–17, 2013

Social Ontology and Social Cognition*

Patrizio Lo Presti

Department of Philosophy and Cognitive Science, Lund University, Sweden
Patrizio.Lo_Presti@fil.lu.se

Abstract

The aim of this paper is to show that there is a reciprocal dependency relationship between social cognition and social ontology. It is argued that, on the one hand, the existence conditions of socially meaningful objects and of social groups are about sucjets' social cognitive processes and interactive patterns and, on the other hand, social cognitive processes and interactive patterns are modulated by socially meaningful objects and social groups. I proceed from a historically informed distinction between social ontologies – between what might be called constructivist and emergentist theories of social reality. I then distinguish three theories of social cognition, theory-theory, simulation theory, and interaction theory, and argue that the first distinction and the latter map onto each other. Finally I argue that the reciprocal dependency between social ontology and social cognition can be justifiably though of as causal in Di Paolo et. al.'s (2010) sense of "downward" or "circular" causation. It is concluded that the dependency between social ontology and social cognition pertain to both a methodological and a phenomenal level. First, research on social ontology depends on research on social cognition; and, secondly, social phenomena, involving socially meaningful objects and groups, influence social cognitive processes and interaction, which in turn influence social phenomena.

1 Introduction: The construction and emergence of the social

How can the contingent empirical fact that we live in a world of nations, cultures, religions, families, and other forms of social relationships that seemingly have causal efficacy on each individual's life, be accommodated with the reductionist realist paradigm prevalent today? As John Searle (2006, p. 13) put it, how, in a world constituted by particles in fields of force, can it be that some carbon based organism after 5 billion years of evolution have created a world of money, property, and government?

These questions form the core of the subject matter of social ontology, a discipline that since Searle's *The Construction of Social Reality* (1995) has surged analytical philosophy and given rise to lively debates. Social ontologists are concerned with the existence conditions of *social phenomena*. Social phenomena are phenomena involving subjects and social relations which 'give rise to' families, groups, organizations, nations, and so on, or units of agency with concomitant roles, rules, norms, and functions.

This paper focuses on the 'give rise to'-relation between subjects and social phenomena. To that end, as an introduction, it is informative to put the 'give rise to'-relation in historical perspective.

*This research was carried out as part of the NormCon project, funded by the European Science Foundation's EUROCORES program EuroUnderstanding, *Understanding and Misunderstanding: Cognition, Communication and Culture*.

Social ontology peaks a longstanding research program that is by and large neglected in todays analytical theorizing. Most analytical philosophers conceive of social phenomena as products of a certain kind. They usually proceed by analyzing the 'give rise to'-relation in terms of the mental states or speech acts of individuals. Social phenomena, in this tradition, are products of the agency of a multiplicity of individuals because, roughly, individuals either knowingly instantiate type-identical mental states or because individuals together declare that social relations or objects have a certain meaning. Thus social phenomena are *constructed*, either out of mental components – beliefs, intentions, desires – the contents of which are shared by individuals who believe that they so share (Tuomela 2003, 2007; Bratman 2009; Gilbert 2006) – or out of speech act components – declaratives, performatives – the utterance of which create social phenomena – families, organizations, nations, money, and so on – if people accept the declarations (Searle 1995, 2006, 2010). Social phenomena, involving socially meaningful objects and relationships, are in the contemporary eye, then, socially constructed. However, this focus on social constructivism in social ontology is itself only the tip of a historical iceberg.

By the turn of the last century there was another conception of the social reality. Emile Durkheim (1895/1972, p. 69) wrote, "Whenever any elements combine and, by the fact of their combination produce new phenomena, it is evident that these phenomena are not given in the elements" According to Durkheim, social and collective phenomena, society at large, are emergent phenomena. Society is a *product of activities* of people but it is no more reducible to individual mental states than life is reducible to mitosis. Georg Simmel had similar ideas about social phenomena when (1910/1971, p. 134) he wrote that people *play society*. That is, according to Simmel as I understand him, social phenomena emerge at the junction of social interactions, and society at large emerges from the social interactions of people who 'play' different roles. I will call this conception of social reality 'emergentist' on the basis that its proponents conceive of social reality as emergent from activity, rather than necessarily constructed intentionally out of speech acts or believed sharing of mental states in the constructivist sense of the previous paragraph.

The above short exercise in the history of social ontology serves to distinguish two alternative accounts of social reality: the *constructivist* and the *emergentist* accounts. Constructivism is the view that people together, through believed sharing of mental states – intentions, goals, commitments, and so on – or through declarative or performative speech acts, *create* social phenomena; social reality is relative to the mental states of individuals aimed at the construction of social relations and objects. This is the view found in much of today's analytically informed social ontology. We find this view in Tuomela's (2007) approach according to which, the existence conditions of social groups, for instance, involve that subjects believe about each other that they believe they are forming a social group. Thus people whose mental states are appropriately related are in position to *create* a social group, or a unit of activity, intentionality, and in general a 'we' of cognition and action. We find constructivism in Gilbert's (1989) semantics of the first person plural pronoun 'we', according to which the referent of 'we' is subjects who have expressed willingness to form a 'we' under conditions of common knowledge. According to Gilbert it is through expressions of appropriate kind under appropriate conditions that collectives, what she calls 'plural subjects', are created. Consider also Searle's (1995, 2010) view that something is a social object only if it is declared to have a function beyond its physical features. On Searle's account, the social functions of objects and persons are relative to the intentionality and recognition of individuals that objects have the functions in question; people must recognize, for instance, that "we accept that these pieces of paper count as money and give their owner the right to buy stuff", and the pieces of paper must be declared to have

that status, in order for the pieces of paper to count as money.[1] In contrast, on the emergentist view, social phenomena are acted out and emerge from social interactions, and are irreducible to mental state or speech act components. From this perspective there are no necessary conditions about acceptance of speech act contents, collective mental states, or believed sharing of mental states, for the emergence of social objects or phenomena. Rather, social reality is the product of patterns of interaction in the sense of being irreducible to individuals' activities or cognitive states and processes; social reality arises from interaction.

So far I have spoken loosely about 'social objects', 'social phenomena', and 'social units'. To clarify, I take social *objects* to be the set of particulars whose existence depends in the constructivist or emergentist sense on social interaction and cognition. Thus rabbit pelt, to use an example of Tuomela's, is not a social object, but when used under the right conditions, e. g., as medium of exchange in social interaction, then it is a social object. If patterns of social interaction recur in which a rabbit pelt is treated as an exchange medium without there being any point at which a rabbit pelt is declared to qualify as a medium of exchange or if there is no believed sharing of acceptance that a rabbit pelt is a medium of exchange, then we can say that the social object, the rabbit pelt as medium of exchange, is an *emergent* social object. In contrast, if declarations or believed sharing of mental states about rabbit pelts is what causes rabbit pelts to count as money, then we can say that the social object, the rabbit pelt as an medium of exchange, is *created* in the constructivist sense. Now, a social *unit* I take to involve similar genealogical processes as social objects, but 'units' or 'unities' relate to social relations rather than to objects. Thus, a family is a social unit, as is a subculture, and in general every instance of a social relation where subjects involved act or cognize as a 'we'. That the genealogy of social units is similar to that of social objects means that just like a rabbit pelt can become a medium of exchange through repeated patterns of social interaction or through declaration or sharing of mental states of acceptance or belief, so a family or a gang can be created through matrimony or vows of allegiance, or emergent in recurring interactions where people, as Simmel would have it, play their respective roles. Lastly, social *phenomena* I take to be phenomena involving social units and objects. Thus it is a social phenomenon that the euro is the medium of exchange in some European countries and that women usually do more household work than men, where money and households are social objects and Europe, and women and men as social groups, are social units. That women's salary is generally lower than men's is a social phenomenon, which if true, is a social fact.

In summary, there are two conceptions of the 'give rise to'-relation between subjects and social phenomena. On the one hand, one can conceive of social phenomena as grounded in declarations or agreement among individuals that certain objects, persons, and relationships are to have certain social statuses, or one can conceive of them as emergent from recurrent interactive patterns not necessarily involving such declarations and agreement.

In the next section, we will see that it is central for any attempt to understand the 'give rise to'-relation between subjects and social phenomena, that is, to understand theories in social ontology, to also understand what the theories presuppose with regard to underlying cognitive states and processes of subjects involved in social phenomena. To provide that understanding I now turn my focus on social cognition.

[1]Constructivism is a widely held approach in social ontology research, and recounting most or even many of its proponents requires too much space. But see, for instance, Gilbert (1990, 2000, 2009), Bratman (1992, 1993), Searle (2006, 2007), Schweikard and Pettit (2006), List and Pettit (2011), Tuomela (2002, 2003).

2 Cognitivism and non-cognitivism

Research on social cognition is research on how people understand others and their social surroundings. More specific, research in social cognition is concerned with cognitive processes that enable subjects to make sense of interaction in social arenas – such as churches, banks – and to act in accordance with how one is meant to act in social arenas. In this section, I will examine three theories of social cognition that can be divided into cognitivist and non-cognitivist theories. My aim here is not adjudication between, but clarification of, these theories with the aim to show how they map onto aforementioned distinctions between constructivism and emergentism in social ontology.

Since the effectuation of false-belief tests (Perner and Wimmer 1983), which are taken to show that children understand that mental states of others can deviate from their own once they have acquired folk-psychological concepts such as 'belief' and 'desire', what has been called theory-theory (TT) has drawn many adherents (Baron-Cohen 1995). According to TT, intersubjective understanding is backed up by mental state-attribution justified by recognition that in certain arenas or in certain interactions people usually have beliefs, intentions, desires, and so on, with a certain content. For instance, when someone reaches for the cookie jar he or she usually has a desire for cookies, a belief that there are cookies in the jar, and an intention to take a cookie from the jar. According to TT, understanding the other person, arriving at the meaning of his or her movement, essentially involves knowing something about cookies and cookie jars and from these premises inferring the other person's intention in terms of his or her beliefs and desires. Inferring intentions in terms of beliefs and desires presupposes having folk-psychological concepts signifying mental states. Since the process is described in terms of inferences, this theory of social cognition is that subjects form a theory about the other person's mental states. Baron Cohen (1995, pp. 3–4) writes, "it is hard for us to make sense of behavior in any other way than via the mentalistic … framework … [A]ttribution of mental states … is our natural way of understanding the social environment" (cf. Toby and Cosmides 1995).

In contrast to TT, and in the wake of neuroscientific research on brain areas functioning as so called 'mirroring' or 'resonance' systems (Gallese and Goldman 1998; Gallese 2005, 2007), a theory according to which social cognitive processes are simulative processes has been suggested. According to simulation theory (ST), people understand each other and their social world by means of running a simulation 'as if' oneself were the other or were in a similar social situation in which an observed other is situated. According to ST, social cognition is not underlain by subjects' mounting of interpretative or inferential processes with the other's behavior or environmental cues as premises, yielding as conclusion what the other means with his or her action or what socially significant environmental cues signify. Rather, "the state ascribed to the target is ascribed as a result of the attributor's instantiating, undergoing, or experiencing, that very state." (Goldman and Sripada 2005, p. 208). Thus, and although in ST there is the notion of agents ascribing mental states and meanings to others, the form of this ascription processes is subjunctive, 'as if', rather than in an inferential theory-like form (Goldman 2005b; Gallese 2005).

Criticism has been mounted against both TT and ST for their commitment to what has been called the *mentalistic assumption*. The mentalistic assumption is that, "mentalizing, or mindreading, underlies basically all social understanding and interaction" (Michael 2011, p. 561). The term 'mindreading' refers to the ascription of mental states involved in social understanding according to both TT and ST. The worries are that, first, if TT or ST were correct, then one should find in phenomenology a corresponding sensation of the theorizing

or simulative processes. But we seem not to be undergoing such phenomenological states when we understand others or our social world (Gallagher and Zahavi 2008, p. 176). And, secondly, even if one assumes that folk-psychological theories or simulative processes are implicit, it still seems that our conceptualizations of the processes underlying social cognition are misleading. Strictly speaking, "there is no neuronal subjunctive" (Gallagher 2007, p. 361). That is, if sub-personal processes of simulation or theorizing involving pretense, 'as if' states, instrumental for mindreading are to function as the explananda of social cognition, then we must conceive of those sub-personal processes as *pretending* and *using* information about the other or the social environment to form a model. But 'pretense' and 'use' are personal-level concepts. Therefore, TT and ST cannot be understood as true descriptions of social cognitive processes, neither at a personal explicit or sub-personal implicit level of description.

The alternative account of social cognition that critics propose is called interaction theory (IT). According to IT, mentalizing or mindreading, i.e., ascriptions or even simulations of mental states do not necessarily underlie social cognition. Rather than the third-person observational, theorizing, or simulative stances TT and ST ascribe to social cognizers, social cognition emerges according to IT in second-person interactions (Gallagher 2008b, pp. 164–5). It is interactive processes themselves, with others and social surroundings, that constitute social understanding and, furthermore, give rise to social meaning (p. 167; cf. De Jaegher et. al. 2010). It is not necessary that interactive processes be supplemented by inferences of simulations.

It should be clear why research on social cognition is important for research in social ontology. Since social ontology produces analyses of social facts and properties – analyses of the existence conditions of such things as money, nations, religion, and families – and since the analyses that are on the table analyze such facts and properties in terms of speech acts, sharing of mental states, or interaction, it is obvious that to understand how social facts and properties can exist it is necessary to understand how subjects can understand each others' mental states, speech act, and actions. Thus social cognitive processing characterizes at least one aspect of the 'give rise to'-relation between subjects and social phenomena.

Since my aim is to show that there is a reciprocal dependency relationship between social ontology and social cognition, both methodologically and phenomenally, I will now try to elucidate, in light of preceding two sections, how questions asked in the two domains map onto each other. In the following sections, I will argue that not only are questions in the two domains linked, but also social phenomena and social cognitive processes themselves exert influence on each other.

Without trying to settle the issue between theory-theory, simulation-theory, and interaction-theory, I suggest that one can distinguish two main approaches to social cognition: *cognitivism* and *non-cognitivism*. Cognitivist approaches to social cognition are characterized by the mentalistic assumption, that is, by their commitment to the claim that social understanding necessarily involves ascription of mental states to others. Non-cognitivist approaches to social cognition are characterized, negatively, by rejection of the mentalistic assumption and, positively, by the claim that perceptual or interactive processes are sufficient for social understanding. Here perceptual and interactive processes are to be understood as *inherently* sense-making. By inherently sense making I mean that it is not necessary that the processes be supplemented by cognitive processes such as inferences or simulations in order for social understanding to be enabled.

From this distinction we can draw two clarifying conclusions regarding commitments of theories in social ontology. First, if communication necessarily involves understanding the meaning, intentions, and communicative intentions of speakers, and if sharing of mental states

necessarily involves mental state-ascriptions as a result of simulation or inference, then constructivist accounts of social reality presupposes a cognitivist account of social cognition. Constructivist accounts of social reality presuppose a cognitivist account of social cognition since Searle (1995, 2006, 2010), whose social ontology is presently one of the most influential, bases his theory on declarative speech acts. And Searle clearly states (1983, p. 166) that, "what one communicates is the content of one's representations", implying that to understand the speech acts with which social objects and units are constructed one must understand and ascribe mental states to speakers. Furthermore, Tuomela, whose social ontology is one of the more prominent amongst those based on believed sharing of mental states (2003, 2007), explicitly states (2007, p. 188) that only collective acceptance that an object has a social meaning and is meant to be used in a certain way can account for the object having that meaning. Second, if it is sufficient for subjects to engage in social interaction that they have intentions the contents of which refer to others, but does not necessarily entail ascription, from theorizing or simulation, of mental states to others, then emergentist accounts of social reality presuppose a non-cognitivist account of social cognition. Emergentist accounts of social reality presuppose a non-cognitivist account of social cognition since Durkheim and Simmel, who I take to be the pioneers of emergentist social ontology (cf. Tollefsen 2002; Gilbert 1989; Greenwood 2003), denied what is now called the mentalistic assumption and claimed that it is people's interactions that constitute social entities. For instance, Simmel (1908/1971, p. 8) wrote, "consciousness of the abstract principle that he is forming society is not present in the individual", suggesting that there need be no (ascription of) beliefs or communicative intentions involved in the emergence of social phenomena.

It is fair to say that the unearthing in this section of the presuppositions of theories in social ontology of theories in social cognition suggests a straightforward mapping of questions asked in the two fields. That is, a constructivist social ontology presupposes that a cognitivist approach to social cognition is supported, whereas an emergentist approach to social ontology does not. The emergentist approach to social ontology is supported by non-cognitivist theories of social cognition. Therefore, constructivist and emergentist social ontology hinge on the plausibility of cognitivist and non-cognitivist theories of social cognition (although a precise forecast for respective social ontologies' ability to handle falsification of theories of social cognition on which they depend cannot at this point be given).

3 Downward causation

The mapping of the socio-cognitive onto the socio-ontological does not simply entail that research on social ontology is aided by research on social cognition. It also entails the reverse relation, that research on social cognition is facilitated by research on social ontology. Furthermore, I will argue in this section that social cognition substantively, not as a research object alone, is facilitated by cognizers being 'situated' in a social reality in a sense to be clarified. It is desirable to first of all investigate the nature of the relation holding between social objects and units and social cognition.

Ezequiel A. Di Paolo and colleagues (2010) use the notion of 'circular' or 'downward' causation: a causal relation holding between emergent entities and low-level processes that give rise to those entities. An emergent entity is described as one "whose characteristics are enabled but not fully determined by the properties of the component processes" (p. 40). This emergent entity in turn "introduces … modulations to the boundary conditions of the lower-level processes that give rise to it" (p. 41). Remember that social phenomena emerge from social cognitive or interactive processes (speech acts, shared mental states, or socially directed

action, depending on what ontology of social reality is preferred). Now, if there is a downward or circular causal relationship between social cognitive processes and social phenomena, that would mean that social cognitive and interactive processes give rise to social phenomena which in turn influence the social cognitive and interactive processes. That is, if social objects and units causally influence people's understanding of each other and their social environments, then that might be understood as social phenomena having downward causal efficacy with regard to the processes of social understanding (or misunderstanding) – the processes from which social phenomena emerge. I will soon illustrate the possibility of the downward, circular relation with two examples. Empirical findings will also be adhered to. But first, let me emphasize that of primary interest for present purposes is vindicating that a downward, circular influence between, on the one hand, social cognitive states and processes and social interactions, and, on the other, social objects and units involved in social phenomena, does obtain. The nature of this relationship is of secondary interest. I will henceforth leave undecided whether this relation is causal in nature and focus instead on the plausibility of the obtaining of the relation.

To illustrate the possibility of downward or circular influence between emergent social phenomena and social cognitive and interactive processes, consider the following example. You're in church with lots of other people filling the rows. The organ is playing and along the aisle two persons are walking solemnly. They stop when they reach the altar and repeat sentences pronounced by the priest. This situation makes sense if you recognize that you are at a wedding. But also, you recognize, or understand, that you are at a wedding by focusing on the social objects and other agents' behavior in this situation. For instance, the altar has a certain meaning, e. g., it is treated as a place where wedding ceremonies, baptisms, and so on, take place. Other objects in the situation have other socially relevant meanings, e. g., the arrangement of benches, peoples' clothing, and so on. Importantly, the set of social objects, units and people in the situation seems to play a central explanatory role in accounting for your, as cognizer, grasping of the social meaning of the situation. But the social meaning of the objects, units, and people is, reflexively, emergent from social cognitive and interactive processes. That is, on the one hand, social phenomena, weddings for instance, emerge from social sense-making and interaction, while, on the other hand, the social phenomen also determine social sense-making and interaction. Social cognition and interaction partly determines the constitution of social reality and the constitution of social reality partly determine social cognitive state and process and social interaction.

Consider an altered version of the wedding example. Suppose that you're giving a lecture at a conference. As you're explaining one of your slides, two persons solemnly dressed as if on a wedding stride towards you between the conference attenders. Something in this situation is terribly wrong, and the obvious reason is that there has been a *misunderstanding*. Why does the appearance of bride and groom *not* make sense? The social environment does match the social interaction; conference halls are standardly not *meant* to house matrimonies, conferences have emerged as gatherings for exchanges of ideas, not for weddings.

The import of these examples is this: objects and persons are *meant* to function, to be used, and to act in certain ways. The existence condition for these functions and roles – these social meanings, the very structure of social environments – is, we have seen, the occurrence of appropriate social cognitive states, processes and interactions of the people involved – their communication, beliefs, or repeated interactions depending on preferred social ontology. What the examples show is that the emergent social objects, roles, functions, and units enable social understanding (or misunderstanding). Understanding and misunderstanding of others and socially relevant objects and events are social cognitive states. Therefore, we can conclude that

social reality, emerging from social cognitive processes and interactions, influence social cognitive processes and interactions and produce states of social understanding (or misunderstanding). Thus social cognitive states and processes and social phenomena circularly influence and perpetuate each other; they form a social circuit, a circular system of reflexive determination.

In the final section before concluding, I will reconnect conclusions drawn about the circular relationship between social cognitive and interactive processes and social phenomena to orthodox contemporary social ontology. The aim will be to find support for my argumentation in some prevalent theories of social ontology. But first, lets summarize our central findings. Two conclusions can be drawn at this point.

First, social cognitive and interactive processes involving objects and persons with social meanings, emergent from social cognitive and interactive processes, can produce social understanding and misunderstanding; i.e., influence social cognition and interaction. Without going into too much detail, this conclusion presupposes that social cognitive processes have access to perceptual, proprioceptive, affective, and other subsystems. This is because identification of social objects and social statuses of others requires access to cues indicating such social meaning – e. g., wedding rings, police badges, uniforms. Also, interaction in accordance with how one is socially meant to interact requires proprioceptive afferent and efferent signals in execution and evaluation of appropriate action. Similarly, affective states, e. g., disapprobation and approval, likely play a role in and are indicative of social understanding and misunderstanding. This does not mean that whenever social understanding or misunderstanding occurs there is some perception of social objects, or proprioceptive or affective state to which explanations of the former necessarily refer. It means that the occurrence or non-occurrence of the latter influence the production of the former. In the next section, I will exemplify how perception can be recruited in social cognition to achieve social understanding or misunderstanding.

Second, the examples considered can be multiplied, and what they show be generalized. Thus take any situation involving social objects or persons with social meaning and shift between introducing and removing them. The prediction is that subjects in the situations imagined will be further removed from or closer to being able to make sense of others and their social surroundings. Since virtually all agency and cognition is agency and cognition situated in a socially meaningful world, questions posed and answered with regard to social agency and cognition and social ontology seem inexorably linked. So there is not only a *methodological* advantage for research in social ontology to be sensitive to research in social cognition, and vice versa, it is also predicted that, *phenomenally*, social cognition and action is sensitive to the ontology of the social environments in which they occur.

I hope to have clarified a sense in which social reality, on the one hand, and social cognition and agency, on the other, partly codetermine each other in a circular manner.

4 Social cognition and social interaction

I said above that the conclusion that social reality influences social cognitive and interactive processes presupposes perceptual access to socially meaningful objects, events, and persons. In this final section, I want to pursue the implications of this presupposition. It will appear that there is support in contemporary orthodox social ontology of my claim that perception of social objects influences social understanding and social interaction.

Searle (1995, p. 85) writes, "we have status indicators in the form of marriage certificates, wedding rings, and title deeds" which serve 'epistemic functions' (p. 120). For Searle, social reality is an epistemically objective, even if ontologically subjective, reality (pp. 8–12). This means that whereas some entities, for instance stones and trees, are ontologically objective in

the sense that their existence is independent of what anyone thinks about such entities, other entities, for instance dollar bills and marriage, depend for their existence on people's assigning and recognizing a social meaning of pieces of paper and social relationships. The 'epistemic functions' of some objects, for instance wedding rings and title deeds, can be thought of as *indicating* what role or function an object or person carries. This is an interesting line of thought connecting to my argument that socially meaningful objects and persons have downward influence on cognition and agency. Because, if, as Searle claims, what meanings persons and objects are bestowed with are indicated, then perceptual access to such indicators certainly implies access to information about social meaning. To conclude that social entities have an influence on social understanding is close to home.

Consider the example of being Secretary General of the U.N. For Searle, having this status *means* that the person has a range of actions open for him or her – a set of 'powers' (p. 106). Interestingly, if social meaning – or social statuses and functions, in Searle's terminology – is indicated by objects so that subjects are epistemically justified in identifying social meaning when perceiving the indicating objects, and if social meanings entail an appropriate way of interacting, then subjects with perceptual access to indicators are in position to make sense of others and their social surroundings. Searle seems never to have seen or have been interested in this implication.

But is it justified to claim that social objects that exert influence, through perception, on social cognitive and interactive processes? Outside of research in social ontology, empirical support for that claim can be found. Shaun Gallagher (2008a) argues that perception of socially meaningful objects is 'smart'. By smart Gallagher means that perception need not be supplemented by other cognitive processes for an observer to make sense of perceptual input (pp. 539–40). Perception of socially relevant objects is 'direct', according to Gallagher, in the sense that there need not be inferential steps or simulative processes premised on perceptual input; perceptual input is in itself sufficiently informative for recognition of something as a car, rather than recognition of something as a car being an inference from perceptual input of metallic mass in a certain shape.

Gallagher have developed the direct perception account, introduced by J. J. Gibson, in the last decade (Gallagher 2001, 2004, 2007, 2008a, 2008b). Although answering how precisely perception can be direct is not a *sine qua non* for my argument to go through, I will give a review of experiments carried out by Marcel (1992), reported by Gallagher and Marcel (1999), about how focus on, and agency in, socially significant situations enhance cognition and agency. Reviewing the experiments is only a way of showing that socially significant objects and events can influence social cognition through perception, and thus that the downward or circular relation between social reality and social cognition and interaction is empirically supported.

Marcel (1992) distinguished three levels of intention formation: intentions in *abstract decontextualized*, in *pragmatically contextualized*, and in *socially contextualized* agency. Abstract decontextualized agency is "detached from what would ordinarily be considered a significant context" (Gallagher and Marcel 1999, p. 9), for instance handling a cylinder shaped object in an experimental setting. Pragmatically contextualized agency is "performed in the course of a natural activity whose purpose arises from personal projects and concerns" (*ibid.*), for instance dishing a teacup. Socially contextualized agency "has a meaning defined by cultural categorizations … and represent states of the self in regard to others" (*ibid.*), for instance serving friends cups of tea at a tea party. What Marcel found was that patients suffering from ideomotor apraxia, that is, persons with difficulty in executing intentions in body movement, had near normal abilities in socially contextualized agency, whereas they had great difficulties

in abstractly decontextualized settings. This led Gallagher and Marcel to conclude that when subjects' intentions are guided by focus on socially significant objects and events involving their social relations to other people their cognitive and agentive performance is enhanced (p. 12; cf. Leontiev and Zaporozhet 1960). Hence there is experimental data in support of the claim that socially meaningful objects and persons influence social cognition and interaction. Since most, if not all, theories in social ontology agree that social objects and meaning emerge from or are constructed in social cognitive and interactive processes, we can conclude that there is empirical support for the claim that there is downward or circular influence between socially meaningful objects and persons, on the one hand, and the social cognitive and interactive processes from which social meaning emerge, on the other.

Reconsidering Searle's notion of social statuses being indicated, it seems we have found in experiments on social cognition a basis for the conclusion of my argument. That is, the social ontology of situations in which people act and cognize influence social cognition and interaction, while social cognition and interaction influence the social ontology of situations in which people act and cognize.

5 Conclusions

Social ontology is about the existence conditions of social phenomena; phenomena involving two or more subjects, their relations and interactions, and often socially meaningful objects involved in interaction. Social cognition is about the sense-making processes, interactive, perceptual, simulative, or theory-like, that enable subjects to understand social phenomena.

In the first sections of this paper, we have seen that social reality is given rise to by social cognitive and interactive processes. The 'give rise to'-relation from such processes to social phenomena can be characterized in several ways. Social objects or units involved in social phenomena can be *created* through speech acts or believed sharing of mental states accepted by a group as the group's goals, beliefs, and so on. Social phenomena can also *emerge* from repeated social interactions in absence of any declarative or performative speech-acts, or believed shared acceptances of goals, beliefs, and so on. The former, constructivist, sense implies some ascription of mental states among people involved in the creation, what in social cognition is called 'mindreading', whereas the latter implies a history of recurrent social interactions not necessarily involving mindreading. The upshot of these implications is methodological: they suggest that research in the domains of social ontology and social cognition is inexorably linked – theoretical presuppositions in one domain are depends on results in the other. A prognosis and desideratum of the state of debate in social ontology and social cognition respectively is, therefore, that only an account that consistently and coherently integrates creation and understanding of social meaning of objects and persons will and should lead the way for future research.

In the latter sections, it has become clear that beyond the desideratum that researchers on social ontology and cognition coordinate efforts, and beyond the prediction that such coordination is fruitful for future research, there is a real reciprocal dependency between social cognitive and interactive processes, on the one hand, and socially meaningful objects and persons, on the other. Thus, a second prognosis and desideratum provided by this paper is that only accounts of social ontology and cognition providing explanation and prediction of social phenomena's and social cognitive processes' reflexive influence on each other will and should lead the way for future research.

In conclusion, whether we analyze the 'give rise to'-relation from subjects to social phenomena in constructivist or emergentist terms, the proposition that I have argued in favor of

suggests itself: social reality is partly determined by underlying social cognitive and interactive processes, and social cognitive and interactive processes are in turn partly determined by the structure of social reality. The result is that the constitution of social reality and the progression of social understanding and interaction can be understood as co-dependent and co-determining.

Acknowledgements: I would like to thank two anonymous referees at Abstracta for crucial insights and criticisms.

References

Baron-Cohen, S. (1995) *Mindblindness: An essay on autism and theory of mind*, Cambridge: MIT Press.

Bratman, M. (1992) 'Shared Cooperative Activity', *The Philosophical Review* 101, 327–341.

Bratman, M. (1993) 'Shared Intentions', *Ethics* 104, 97–113.

Bratman, M. (1999) *Faces of Intention: Selected Essays on Intention and Agency*, Cambridge: Cambridge University Press.

Bratman, M. (2009) 'Modest sociality and the distinctiveness of intention', *Philosophical Studies* 144: 149–165.

Butterfill, S. (in press) 'Interacting mindreaders', *Philosophical Studies*.

Davidson, D. (2001) 'Agency', In D. Davidson (ed.) *Essays on Actions and Events*, Oxford: Oxford University Press.

De Jaegher, H. (2009) 'Social understanding through direct perception? Yes, by interacting', *Consciousness and Cognition* 18, 535–542.

De Jaegher, H., Di Paolo, E.A. & Gallagher, S. (2010) 'Can social interaction constitute social cognition?', *Trends in Cognitive Science* 14, 441–447.

Di Paolo, E.A., Rohde, M. & De Jaegher, H. (2010) 'Horizons for the Enactive Mind: Values, Social Interaction, and Play', In J. Stewart, O. Gapenne and E.A. Di Paolo (eds.) *Enaction: Toward a New Paradigm for Cognitive Science*, Cambridge: MIT Press.

Durkheim, E. (1950) 'Les règles de la methode sociologique', In A. Giddens (ed.) *Emile Durkheim: Selected Writings* (1972), New York: Cambridge University Press.

Gallagher, S. (2001) 'The practice of mind: theory, simulation or primary interaction?', *Journal of Consciousness Studies* 8, 83–108.

Gallagher, S. (2004) 'Situational Understanding: A Gurwitschian Critique of Theory of Mind', In L. Embree (ed.) *Gurwitsch's Relevancy for Cognitive Science*, Dodrecht: Springer.

Gallagher, S. (2007) 'Simulation trouble', *Social Neuroscience* 2, 353–365.

Gallagher, S. (2008a) 'Direct perception in the intersubjective context', *Consciousness and Cognition* 17, 535–543.

Gallagher, S. (2008b) 'Inference or interaction: social cognition without precursors', *Philosophical Explorations* 11, 163–174.

Gallagher, S. & Marcel, A.J. (1999) 'The Self in Contextualized Action', *Journal of Consciousness Studies* 6, 4–30.

Gallagher, S. & Zahavi, D. (2008) *The phenomenological mind*, London: Routledge.

Gallese, V. (2005) "Being like me': Self-other identity, mirror neurons and empathy', In S. Hurley and N. Chater (eds.), *Perspectives on imitation* vol. 1, Cambridge: MIT Press.

Gallese, V. (2007) 'Before and below 'theory of mind': Embodied simulation and the neural correlates of social cognition', *Philosophical Transactions of the Royal Society B* 362, 659–669.

Gallese, V. & Goldman, A.I. (1998) 'Mirror neurons and the simulation theory of mind-reading', *Trends in Cognitive Science* 2, 493–501.

Greenwood, J.D. (2003) 'Social Facts, Social Groups and Social Explanation', *Nôus* 37, 93–112.

Goldman, A.I. (2005) 'Imitation, Mind Reading, and Simulation', In S. Hurley and N. Chater (eds.), *Perspectives on imitation* vol. 2, Cambridge: MIT Press.

Goldman, A.I. and Sripada, C.S. (2005) 'Simulationist models of face-based emotion recognition', *Cognition* 94, 193–213.

Grice, H.P. (1989) *Studies in the Way of Words*, Harvard: Harvard University Press.

Gilbert, M. (1989) *On Social Facts*, Princeton: Princeton University Press.

Gilbert, M. (1990) 'Walking Together: A Paradigmatic Social Phenomenon', *Midwest Studies in Philosophy* 15: 1–14.

Gilbert, M. (2006) 'Rationality in Collective Action', *Philosophy of the Social Sciences* 36: 3–17.

Gilbert, M. (2009) 'Shared intention and personal intentions', *Philosophical Studies* 144: 167–187.

Lavelle, J.S. (2012) 'Theory-Theory and the Direct Perception of Mental States', *Review of Philosophy and Psychology* 3, 213–230.

Leontiev, A.N. & Zaporozhet, A.V. (1960) *Recovery of Hand Function*, London: Pergamon.

List, C. & Petitt, P. (2011) *Group Agency: the possibility, design, and status of corporate agents*, Oxford: Oxford University Press.

Marcel, A.J. (1992) 'The personal level in cognitive rehabilitation', In N. Von Steinbüchel, E. Pöppel and D. Cramon (eds.) *Neurophysiological Rehabilitation*, Berlin: Springer.

Michael, J. (2011) 'Interactionism and Mindreading', *Review of Philosophy and Psychology* 2, 559–578.

Miller, K. & Tuomela, R. (1988) 'We-Intentions', *Philosophical Studies* 53, 367–389.

Scweikard, D. & Pettit, P. (2006) 'Joint Actions and Group Agents', *Philosophy of the Social Sciences* 36, 18–39.

Searle, J. (1969) *A Theory of Speech Acts: An Essay in the Philosophy of Language*, Cambridge: Cambridge University Press.

Searle, J. (1983) *Intentionality: An Essay in the Philosophy of Mind*, Cambridge: Cambridge University Press.

Searle, J. (1995) *The Construction of Social Reality*, New York: The Free Press.

Searle, J. (2003) 'Social Ontology and Political Power', In F.F. Schmitt (ed.), *Socializing Metaphysics: The Nature of Social Reality*, Oxford: Rowman & Littlefield.

Searle, J. (2006) 'Social Ontology: Some Basic Principles', *Anthropological Theory* 6, 12–29.

Searle, J. (2010) *Making The Social World: The Structure of Human Civilization*, Oxford: Oxford University Press.

Simmel, G. (1908) 'Exkurs über das Problem: Wie ist Gesellschaft möglich?', In D.N. Levine (ed.) *George Simmel: On individuality and social forms* (1971), Chicago: Chicago University Press.

Toby, J. & Cosmides, L. (1995) 'Foreword', In S. Baron-Cohen (author) *Mindblindness: An essay on autism and theory of mind*, Cambridge: MIT Press.

Tollefsen, D.P. (2002) 'Collective Intentionality and the Social Sciences', *Philosophy of the Social Sciences* 32, 25–50.

Tuomela, R. (2002) *The Philosophy of Social Practives: A Collective Acceptance View*, Cambridge: Cambridge University Press.

Tuomela, R. (2003) 'Collective Acceptance, Social Institutions, and Social Reality', *American Journal of Economics and Sociology* 62, 123–165.

Tuomela, R. (2005) 'We-Intentions Revisited', *Philosophical Studies* 125, 327–369.

Tuomela, R. (2006) 'Joint Intention, We-Mode and I-Mode', *Midwest Studies in Philosophy* 30, 35–58.

Tuomela, R. (2007) *The Philosophy of Sociality: The Shared Point of View*, Oxford: Oxford University Press.

Tuomela, R. & M. Tuomela (2005) 'Cooperation and trust in group context', *Mind & Society* 4, 49–84.

Wimmer, H. & Perner, J. (1983) 'Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception', *Cognition* 13, 103–128.

Paper 3

HUMAN AFFAIRS *25*, 110–120, 2015 DOI: 10.1515/humaff-2015-0009

RULE-FOLLOWING, MEANING CONSTITUTION, AND ENACTION

PATRIZIO LO PRESTI

Abstract: The paper submits a criticism of the standard formulation of Wittgenstein's rule-following paradox. According to the standard formulation, influenced by Kripke, the paradox invites us to consider what mental or behavioral items could constitute meaning. The author proposes instead an enactivist understanding of the paradox. On this account there is no essential gap between mental items and behavioral patterns such that the paradox enforces a choice between meaning being constituted either internally 'in mind,' or externally 'in behavior.' The paper begins with an introduction to the paradox and then presents arguments against standard solutions. It ends with the enactivist proposal, admitting that although much more needs to be said before it can be established as a full-fledged alternative, it nonetheless holds some promise both for revising our understanding of the paradox and for the formulation of a novel solution.
Keywords: Wittgenstein; rule-following; meaning; enaction; interaction.

The rule-following paradox (Wittgenstein, 1953) is arguably the most "radical" problem for modern philosophy (Kripke, 1982, p. 1), its "iconoclastic" consequences unmatched (Pettit, 2002, p. 31). Allegedly at stake are the foundations of mathematics (Wittgenstein, 1956), rules of logic, and meaningful language. The paradox has even been called an "antinomy of pure reason" (Boghossian, 2012, p. 47), and a modern "scandal of philosophy" (Peacocke, 2012, p. 66).

Here is the passage in which Wittgenstein formulates the paradox:

> This was our paradox: no course of action could be determined by a rule, because every course of action can be made out to accord with the rule. The answer was: if everything can be made out to accord with the rule, then it can also be made out to conflict with it. And so there would be neither accord nor conflict here (Wittgenstein, 1953, § 201).

The reasoning preceding the quoted section concerns language meaning. The question is: how can symbols have definite meanings such that in using one now I can (fail to) use it in accord with that meaning? Wittgenstein notices that token symbol uses initiate many paths for future use, which can be made out to accord or conflict with past uses. So, there is in principle an indefinite number of interpretations of symbols or words that are all within the set that can be made out to accord and conflict with past use. Hence, no interpretation of past

use can constitute a rule that singles out one definite meaning. Hence the paradox: while we use words in definite ways and do so quite successfully to communicate in everyday practice, there seems to be no rule we may rely on as a determiner of correct use.

Wittgenstein responds to these considerations in the second half of the quoted section:

> It can be seen that there is a misunderstanding here from the mere fact that in the course of our argument we give one interpretation after another; as if each one contended us at least for a moment, until we thought of yet another standing behind it. What this shows is that there is a way of grasping a rule which is *not* an *interpretation*, but which is exhibited in what we call "obeying the rule" and "going against it" in actual cases (Wittgenstein, 1953, § 201).

Perhaps, then, we might better understand what constitutes meaning by abandoning the idea that it is constituted by interpretation. We then avoid the problematic property of interpretations that they are themselves correct or incorrect with respect to what they interpret (Wittgenstein, 1953, § 198), according to a rule—a property that threatens to gear our investigations towards a regress of interpretations. The question then is: with what do we replace interpretations as constitutive of meaning? How are what we call 'obeying' and 'going against' the rule exhibited in actual cases? In the closing three lines of the quoted passage Wittgenstein makes an inference that might hold the clue:

> Hence there is an inclination to say: every action according to the rule is an interpretation. But we ought to restrict the term "interpretation" to the substitution of one expression of the rule for another (Wittgenstein, 1953, § 201).

It is the tendency to interpret 'interpretation' as a mental act that constitutes meaning that spells trouble. Thus Wittgenstein urges us to limit 'interpretation' to substitutions of one expression of rule. Rule-expressions include *doing* that which accords with or violates customs and practices established over time (Wittgenstein, 1953, §§ 199, 217, 219). Again, what is correct or incorrect with respect to custom behavior, linguistic or otherwise, ought not to be thought of as behavior that accords or conflicts with a mental act of interpretation, lest we fall back into the original dilemma. Rather we ought to think of 'obeying' and 'going against' practice and customs as 'exhibiting' actual case-by-case agreement on meaning.

But if being in accord with or in violation of a rule does not consist in an interpretation in the sense of mental acts that distinguish accord from conflict, then the question becomes: what makes the everyday engagement and exhibiting of a rule different and insulated from the worry that, really, there are no rules? Kripke (1982, pp. 86-87) holds that since no mental act or behavioral disposition could constitute rule-following we have to abandon the idea of a "straight" solution to the rule-following paradox. A straight solution would be one according to which there are necessary and sufficient conditions for saying that a behavior is a case of rule-following. Kripke presents the problem of rule-following as a choice between the Scylla of rule-following consisting in mental acts or intentions, on the one hand, and the Charybdis of rule-following consisting in behavioral dispositions on the other. He writes, concerning the addition function:

> if the sceptic is right, the concepts of meaning and of intending one function rather than another will make no sense. For the sceptic holds that no fact about my past history—nothing

> that was ever in my mind, or in my external behavior—establishes that I meant plus rather than quus … But if this is correct, there can of course be no fact about which function I meant, and if there can be no fact about which particular function I meant in the *past*, there can be none in the *present* either (Kripke, 1982, p. 13).

Abstracting from this particular example, we clearly see the two horns of a dilemma. The first is to suppose that something 'internal to mind' constitutes meaning. The other is to suppose that something in 'external behavior' constitutes meaning—presumably, 'external' in the sense of what is left of what I do when I have abstracted from my doings the intentions that premise the action. These two options are what Kripke calls "straight" solutions to the paradox. He rejects both in favor of a "skeptical" solution; one in which no *fact* constitutes rule-following (Kripke, 1982, pp. 86-87). In the following sections (1 and 2) I will face up to the two horns in order to expose any promise or further threat they may present. In the end I accept Kripke's rejection of both. But I do not accept that the rejection supports skepticism (sections 3 and 4).

My aim is to thread a third non-skeptical route. The third route is enactivist in spirit. It holds that behavior does not merely express-exhibit meaning but, in a sense to be specified, meaning is *co-constituted* by mental acts and embodied behavior of interpretation and representation of regularities of one's self and others, thereby giving rise to an *operationally closed* dynamic and *circular causal* pattern encompassing bodily and mental action that reciprocally affords future bodily and mental action, thus perpetuating a domain of meaning.

A caveat: the enactivist proposal is not, as presented in this piece, able to answer all the intricacies of meaning constitution. This piece submits a dilemma, presents why the suggested solution has not been found persuasive and suggests an alternative, non-skeptical understanding of the dilemma. As such it is concerned with theory development and not with defending an established paradigm. The reader is also advised that purely exegetical points are not made or intended. This is not a prescriptive account of how a text, in particular one by Wittgenstein, ought to be read. Objections that pertain to exegesis alone should therefore, unless relevant to the validity of the argument, be put aside for present purposes.

Dispositions—the first horn

A natural reaction to Wittgenstein's idea that 'just acting' expresses-exhibits a rule (Wittgenstein, 1953, §§ 217, 219) is to argue that behavioral *dispositions* constitute meaning. The false impression that rules determine meaning may then be refined through an identification of the patterns people are disposed to follow. These dispositions can be stated in generalized form *as if* rules. What constitutes following a rule is then one's being in accord with dispositions that constitute meaning. So we would have a natural reduction of rules to behavioral dispositions that in turn constitute meaning.

Varieties of dispositionalism (e.g., Pettit, 2002; cf. Millikan, 1990) have had their share of criticism in recent debate. On the face of it, dispositionalism might seem the most natural response to our dilemma, which is perhaps why Kripke (1982, pp. 22-32) devoted considerable space in an attempt to refute it. Today dispositionalism appears less attractive. In this section I present some arguments against dispositionalism.

Kripke pointed out that dispositions that subjects embody encounter the 'problem of finiteness' (*ibid*., p. 26). The problem is that past instances of practice coincide with indefinite ways of proceeding and none of them enjoy the privilege of being exclusively true to past practice. In multiplication, for example, the multiplicand may be too large for it to be possible for it to be read off from subjective dispositions about which product one is disposed to getting. In that sense, although one might *actually* be disposed to multiplying one way rather than another, dispositions cannot *constitute* the meaning of 'multiplication.' Subjects might be disposed to behaving in one particular way rather than another, even where there are in principle an infinite number of future instances. But that does not give us any reason to assume that a subject's dispositions will not change over time as he re-engages that practice, or that, when he engages it at any instant, that the dispositions operative at *that* instant determine the correctness of conditions for future engagement. Due to the finitude of subjective dispositions we are in the dark about whether dispositions can stay true to, or at any moment constitute, the meaning of a practice, not just at the present moment but over time as well (Kripke, 1982, p. 27; cf. Millikan, 1990, p. 327).

Another influential objection to a dispositionalist account of meaning constitution is that people are disposed to making *mistakes* (Kripke, 1982, pp. 28-32). If meaning is to be read-off from people's dispositions, then there is no reason in principle to deny, for instance, that 'multiplication' means something different in conditions of fatigue when people are disposed to getting 2 as the product from 7 and 9 or are disposed to clinching the person they mean to 'converse' with rather than 'wrestle'. It has been argued that meaning constituting dispositions should be identified as those, which, in the 'right' circumstances, pick out a definite meaning. The question then is: which circumstances are 'right'? It has been suggested that the 'right' conditions are those in which 'biological competences' operate without interference (Millikan, 1990), and as those that result in interpersonal and intrapersonal-intertemporal congruence in understandings of the meaning of a practice (Pettit, 1990).

I cannot, for lack of space, provide detailed criticism of proposals as to what the 'right' conditions are for dispositions to operate so as to constitute meaning—nor can I assess the many dispositionalist counterarguments (Soames, 1998; Horwich, 1998). But we may consider three reasons why disposition-firings, even in the 'right' circumstances, cannot be the meaning constituting facts we are looking for.

One problem is that idealizing conditions, the 'right' conditions, tends to *trivialize* matters. Suppose we could delimit the conditions that were right. We would be left with a trivial claim to the effect, for example, that Jones means *horse* by 'horse' if and only if Jones applies 'horse' to *horses* in the right conditions. For instance, if on dark nights Jones is disposed to mistaking horse-like cows in the distance for horses, 'horse' still means *horse* (rather than *horse or horse-like cow*) because Jones would apply 'horse' only to *horses* if the conditions were right. But, surely, if we could eliminate deceptive counterfactuals it is *trivially true* that Jones *would* apply 'horse' only to *horses*. This analysis does not avail us. We want an answer to the question of what fact constitutes 'horse' meaning *horse*, not *how Jones applies* the word 'horse.' The latter only answers the analytically secondary question: *given* that 'horse' means *horse*, under what conditions will Jones refer successfully? Saying that Jones is disposed to applying 'horse' only to *horses*, in the right conditions, tells us

nothing about what constitutes the meaning of 'horse' to begin with. Unless we know what makes conditions 'right' for dispositions to operate in them, not what people are disposed to doing when the conditions are 'right,' we cannot distinguish which of them are dispositions to making mistakes, and which constitute meaning.

It is tempting to avoid triviality by substantiating 'right' by employing a term δ, say, 'interpersonal convergence on meaning' (cf. Pettit, 1990). But in that case vacuity is traded for either circularity or regress (Boghossian, 2005, pp. 192-93). If the new definition of 'right' conditions does not involve normative terms, then it tells us nothing about why some conditions are 'right' and others are not, and so the initial challenge has not been tackled. On the other hand, if δ is normatively loaded, then we may ask what makes the conditions it denotes any more 'right' than other conditions. In principle it is an open question; if δ denotes something non-normative, are the δ-conditions *right*? No answer to this question can recast δ, without regress, in *non*-normative terms. But if δ denotes something normative, then it is an open question as to why we should accept it. A skeptic might disagree that *those* are the 'right' conditions. If, in response, one falls back on a non-normative notion of δ, the regress reappears. But if, instead, one responds by pointing to some further normative considerations by virtue of which δ-conditions are normatively 'right', then the skeptic will press for an argument as to why the new property thus picked out makes the conditions more 'right' than others. And this dispute can continue ad infinitum.

Now, dispositionalist accounts of meaning constituting facts encounter some serious problems. I am not suggesting that arguments against dispositionalism are conclusive. Only that, at the moment, dispositionalism faces obstacles that are too serious for it to be fruitful.

Intentions—the second horn

Another natural response to the rule-following paradox is to point out that the *interpretations*-regress that Wittgenstein encounters (Wittgenstein, 1953, §§ 201, 217) does not exhaust which mental acts could constitute meaning. This is the second horn of our dilemma.

An obvious candidate for meaning constituting mental acts is *intentions*. Intentions with general content, for instance, 'do A whenever conditions are C,' provide directives that quantify potentially infinite situations. In that sense an intention view may offer a stepping-stone to overcoming the problem that prior practice presents only a finitude of instances coinciding with in principle indefinite number of rules for future practice (Wright, 2001, pp. 125-26). We may say that meaning is constituted by the generalized content of intentions from which people engage in a practice. This, then, is how rule-following is possible: W means M in conditions C if and only if people in C intend to use W to mean M. Thus, if we recognize that conditions are of type C, and if we have an intention with the general content to W to mean M whenever C, then to W in C is to follow the rule for meaning M in C.

But a regress lurks here as well. The problem (cf. Boghossian, 2012) is that the subject will have to track the conditions, C, picked out by the generalized content of her intentions in order to identify what to do in a practice she is currently engaged in or about to engage in. If she intends to follow a rule of the form 'whenever C, do A!' she must believe that she is in C and infer 'do A!' The subject must infer that she is in a condition fitting the generalized content of her intention in order for her to behave according to the rule. But

now, the inference involved in this reasoning itself requires that she follows a rule—the rule of inference from 'if I am in C, I ought to do A!' and 'I am in C' to 'I ought to do A!' When the subject moves from the content of her intention via the premise that she is in C to the conclusion that she ought to do A, she is using a rule of inference in order to follow a rule! As Boghossian puts it:

> If on the Intention view, rule-following always requires inference; and if inference is itself always a form of rule-following, then the Intention view would look to be hopeless: under its terms, following a rule requires embarking upon a vicious regress in which we succeed in following no rule (2012, p. 41).

In order to avoid the regress one could dismiss the premise that inference is a matter of rule-following. We may leave this as an open possibility, but note that giving up the idea that inferences are instances of rule-following is radical indeed. The very idea that, from 'P' and 'if P then Q', concluding 'Q' is permitted, relies on the Modus Ponens rule. If there were no rule allowing this inference, many arguments, in philosophy or otherwise, would stand on shaky ground. If invoking intentions with general content was supposed to bridge the finitude problem, then it appears no more plausible than the idea that dispositions could bridge that gap.

Perhaps, though intentions could be interpreted as *substituting* rules, rather than *constituting* rules, as per Wittgenstein's suggestion above. Rather than intentions constituting something in addition, i.e., rules, which then constitute meaning, people intend to behave in definite ways and those intentions simply *are* what we call 'rules.' In that case, the intention view is not that people intend *to follow rules* but that *intentions* constitute meaning directly. Thus I may confidently say that '×' means *multiplication* because that is how I intend to use '×'.

However, if intentions constitute meaning, then whatever a subject *intends* the meaning of a practice to be will *be* its meaning. We take this route at the peril of collapsing the distinction between behaving in accord with the meaning of a practice, on the one hand, and intending to behave in accord with its meaning, on the other. Jones may intend to mean something by '×' that he didn't mean earlier. Insofar as intentions constitute meaning, though, there is no principled distinction to be made between Jones failing to mean a definite function by '×' and his deciding upon a new meaning that he *now* intends for it. Whatever he intends its meaning to be is its meaning (Wittgenstein, 1953, § 202). If he is wrong in the light of his earlier intentions, he may modify his intentions to get back on track—or, which amounts to the same thing, modify the meaning by modifying his intentions. But then, playing chess, for instance, and intending that what one is doing is playing chess amounts to the same thing—which is absurd. Determining what Jones is doing would be equivalent to determining what he intends to do. Jones may be throwing pebbles in a well intending this to mean 'checkmate' or 'touch down,' and on to the present proposal that would then be what his pebble throwing means.

There are, then, important obstacles at the second horn of the dilemma and, aside from details we cannot do justice to here, this is where the discussion stands. Some opt for the first horn, arguing that behavioral dispositions constitute meaning. Others opt for the second horn, arguing that intentional or other representational states constitute meaning. What remains to be said in defense of a non-skeptical approach to meaning constitution? I devote the final two sections to that question.

Inadequate constraints

Our question was: which facts constitute meaning (Cf. Boghossian, 1989, p. 515)? The rejection of both 'internalist' and 'externalist' options force our answer towards skepticism. I want to argue, on the contrary, that we are not bound to either horn or to skepticism. I will do so via an introduction to *enactivist* philosophy.[1]

A central idea of enactivism is that subjects do not construct detailed internal representation of an external world from which they infer or recover meaning that then informs action (Varela, Thompson, & Rosch, 1991, p. 149). Nor do subjects interpret past practices, in the narrow sense of taking such practices as input, which they process in order to derive meaning and to output behavior (*ibid.*, pp. 156-7). Meaning is neither external and recovered nor internal and projected.

The rejection of meaning as recovered from the external world does not, however, place enactivism on the second horn, on which meaning is constituted internally and then *projected* onto behavioral patterns. Meaning is understood rather as *enacted* (*ibid.*, pp. 149, 173). The verb 'to enact' signifies a temporally extended process of embodied and social interaction in which what subjects do, and the way the world is, co-specify a *relational* domain of viability (Thompson, 2007, p. 74). It isn't from the content of intentional or other representational states, on one side, or from mechanistically described behavior or environmental structures, on the other, that meaning is somehow projected or recovered. Instead, social and physical interactive processes are understood as driven by the dual force of internal and external dynamics that molds a relational domain "in between" the internal and external (De Jaegher & Froese, 2009, p. 447). Meaning consists in saliencies to think and act one way rather than another. These saliences are enacted as people interact with each other and their environments, and are ultimately grounded both in internal and external, biological and (social and physical) interactive processes. What we think of, interpret, or judge as *meaning* (what may appear as a *rule*) is the present state of such unfolding processes. Importantly, enaction of meaning involves the *co*-constitution of the unfolding of practice, involving both internal and external processes, neither of which is analytically primary but on a par.

From this point of view, the first two horns considered in the previous sections are avoided. But importantly, enactivism does not square well with skepticism. Enactivism suggests that we abandon the assumption that the two horns are mutually exclusive and exhaustive of alternative solutions to the rule-following paradox by urging us to not accept the either-or alternatives of internalism and externalism at the very beginning. If enactivism has anything going for it, then the question about constitution cannot be formulated on the assumption that we have already accepted that it raises two horns, represented by the 'either internal-mental or external-behavioral'-disjunction, or forces us towards skepticism. In order to assess the force of this third non-skeptical solution, much more needs to be said about enactivism.

---

[1] Enactivism originated as a movement in the biology of cognition and the organisation of living (Maturana & Varela, 1980). Since its inception it has developed into one of the staunchest adversaries of the rules-and-representations model of mind and cognition (Varela, Thompson, & Rosch, 1991). It is known for its animosity to the supposedly Cartesian separation of mind and body (Thompson, 2007, pp. 226-30). I cannot give a full exposé of enactivism here and refer the reader to the works quoted.

Enacting meaning

Key concepts of enactivism are *autonomy* and *sense-making*. I introduce both here and then elaborate in more detail the enactivist understanding of meaning constitution introduced in the previous section.

Autonomy

Autonomy signifies the following property of agents: they are self-generating entities that through their biochemical processes and interactions with other autonomous agents and environments sustain an identity over time. Metabolism secures that living organisms' organization and structure is kept autonomous over time courtesy of being thermodynamically permeable but operationally closed systems. This allows intake of self-sustaining material that contributes energy to the organism, thus enabling it to *act* as an autonomous being with its own organization and structure. The continuity of autonomy achieved through this biological interactive process "establishes a *perspective* on the world with its own normativity, which is the counterpart of the agent being a center of *activity* in the world" (De Jaegher & Di Paolo, 2007, p. 488). This establishing of a perspective on the world is simultaneously a realization of a domain of *behavioral viability*, or *value*, relative to the organism's biological autonomy and the structure of its environment. Behavioral patterns become valuable as a function of their tendency to perpetuate autonomy in precarious conditions. Consequences of interactive processes "have significance or value … in relation to the processes of its [the agent's] identity generation" (De Jaegher & Froese, 2009, p. 447). That is, interaction becomes inherently having value and significance for autonomous agents that occupy a center of activity in the world, both affectively through experience of tensions between selfhood and alterity, as well as thermodynamically and biologically through the unfolding interaction's impingement on autonomy.

Sense-making

Sense-making is the process in which autonomous agents create and appreciate meaning as consequence of the establishment of value through interaction with others and environments:

> Exchanges with the environment are inherently significant for the cogniser and this is a definitional property of a cognitive system: the creation and appreciation of meaning or *sense-making* for short. … [S]ense-making is an inherently active concept. Organisms do not passively receive information from their environments, which they then translate into internal representations whose significant value is to be added later. Natural cognitive systems are simply not in the business of accessing their world in order to build accurate pictures of it. They actively participate in the generation of meaning in what matters to them; they enact a world (De Jaegher & Di Paolo, 2007, p. 488).

Enactivists reject "the traditional dichotomy between internal and external determinants of behaviour" (De Jaegher & Di Paolo, 2008, p. 2), replacing it with the idea that internal and external dynamics, biological viability and (social and physical) interaction, co-specify conditions in which meaning is enacted (*ibid.*).

Can we, from an enactivist perspective, avoid the 'either-or' dilemma about meaning constituting facts? We can, but at present it is a long shot. We may concede that no instruction or rule that I give myself, no generalized intentional content "engraved on my mind as on a slate" (Kripke, 1982, p. 15), constitutes meaning. Nor is it plausible that meaning is constituted by dispositions or something in the external world, for reasons considered earlier. But this is not a concession that *no fact* (*ibid*., p. 13) constitutes meaning. Skepticism follows only if an abstraction of mind from action is assumed as already accepted. From an enactivist perspective, the abstraction is non-mandatory. The enactivist proposal is to treat mind and autonomous action as *co*-constitutive of meaning, while engaging in practice *reciprocally* specifies paths for future autonomous sense-making. What people have in mind, and what they do (their linguistic, arithmetic, social and embodied action) are not separate domains, but are co-dependent in the enaction of a meaningful world in which agents emerge as *minded* beings; as beings having mental content. Internal biological and metabolic processes together with (social and physical) interaction *co*-constitute practices whose history and progression in turn establish *value* and *concern* for thinking and behaving one way rather than another. Therefore neither internalism nor externalism will be satisfying (Di Paolo, 2009). Neither internal nor external processes are, as such, sufficient for meaning constitution in abstraction from their mutual co-specification of dynamical enactive processes.

It will immediately be observed that there is an enormous gap to account for between the enaction of biological viability, at one end, and mathematical sophistication, at the other. I will return to this shortly.

But first, an important aspect of meaning enaction is that meaning is essentially dynamic. The mental and behavioral aspects of meaning constituting processes are susceptible to the dynamics of internal as well as external (social and physical) processes. Vice versa, the domain of viability for thinking in the world, enacted through the interdependent dynamics of internal and external processes, is in turn susceptible to modulation through the unfolding of those internal and external, biological and interactive (social and physical) processes. Thus the domain of meaning that autonomous sense-making organisms enact can itself be understood as operationally closed at a higher level of description. The enacted world is essentially an *interactive* domain *in between* agents and their environments (De Jaegher et al., 2010) that is *itself* operationally closed. The interactive domain of value and concern is higher-level in that it emerges from lower-level biological, embodied and affective dynamics of agents and environments. It is operationally closed in the sense that it exerts a "downward" causal force (Di Paolo et al., 2010; Lo Presti, 2013, pp. 11-12) on the unfolding of lower-level biological, metabolic and (social and physical) interactive processes. However, the sense-making that is thus being enacted is not *biologically* closed, since it encompasses multi-agent systems and their environments in interaction as well as their histories of interaction and sense-making. In this interactive process agents *co*-contribute to the enaction of a domain of value and concern but are equally susceptible to the causal feedback from that enacted domain. This two-level interdependence approach to meaning constitution renders off-track the idea that generalized rules or directions (in mind or behavior) constitute meaning. The meaning of future practice must be filtered through the dynamics of imminent interaction. Imminent interaction, though its dynamics are somewhat restricted by histories of

interactions and the values and concern previously enacted, involves an embodied-affective negotiation of here-and-now (physical and social) contingencies, and as such has a causal force which thrusts interaction dynamics along slightly, or radically, different trajectories. This causal process then molds the unfolding dynamics, again effectuating slight or more radical changes, or preservation of, future interaction dynamics. Constituting meaning becomes inherently *activistic*.

Wittgenstein's rejection of the metaphor that meaning is determined by a rule whose steps "have already been taken" and which "is to be followed through the whole of space" (Wittgenstein, 1953, § 219) appears to be right. On an enactivist construal, we may say that meaning is enacted in *taking steps* (in language use and calculation; in engagement). The steps that have already been taken specify a way of thinking and acting in the future. People do not engage practices as if they were novel, for each instance with an erased mental, behavioral, and interactive history. We engage practices whose antecedently enacted meaning *affords* ways of thinking, interpreting, perceiving, and interacting; practices are enacted as making sense relative to our concerns. In this process, the co-contribution of mental, behavioral, and (social and physical) interaction is the factual constitutive ground of meaning.

The outcome of this suggestion is an alternative to the either-or story reviewed earlier. What people do and think specifies a meaning that specifies a way for people to think and behave in the future (e.g. in language use or in doing mathematics). We are not, then, required to infer from prior instances what the meaning of a practice is to form a conclusion about what to do or say now or in the future. We *may* do so, but it does not *constitute* meaning. Neither is it necessary to have general instructions or rules 'in mind', or the 'right' dispositions. Meaning, and knowing how to proceed in accord with the meaning of a practice (Wittgenstein, 1953, §§ 151-54), is a question of mental content and behavior being fostered by, as well as fostering, engagement in day-to-day activities. Meaning constitution is activism.

Enactivism, in conclusion, provides one venue for criticizing the very formulation of the rule-following paradox. It suffices for my purposes if the criticism has initial plausibility. If it does, then we can conclude that the two horns and skepticism, developed over the last three decades or so, have, indeed, *blinded* us from a novel way of proceeding in our philosophical investigating on meaning. If, in addition, enactivism is promising, then formulating a fourth venue is within reach. One should expect that enactivism would have to withstand much criticism to emerge as a feasible alternative. Either way, it is no less significant a conclusion that at the moment we seem not to have properly understood our dilemma.[2]

References

Boghossian, P. (1989). The rule-following considerations. *Mind*, *98*(392), 507-549.
Boghossian, P. (2005). Rules, meaning and intention: Discussion. *Philosophical Studies*, *124*, 185-197.

[2] This research was funded by the European Science Foundation's EUROCORES program EuroUnderstanding, and was carried out as part of the NormCon project. I would like to thank the members of that project for helping me develop the argument here presented.

Boghossian, P. (2012). Blind rule-following. In A. Coliva (Ed.), *Mind, meaning, and knowledge: Themes from the philosophy of Crispin Wright*. Oxford: Oxford University Press.

De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making. *Phenomenology and the Cognitive Sciences*, *6*, 485-507.

De Jaegher, H., & Di Paolo, E. (2008). Making sense in participation: An enactive approach to social cognition. In F. Morganti, A. Carassa, & G. Riva (Eds.), *Enacting intersubjectivity: A cognitive and social perspective to the study of interactions*. IOS Press: Amsterdam.

De Jaegher, H., & Froese, T. (2009). On the role of social interaction in individual agency. *Adaptive Behavior*, *17*(5), 444-460.

De Jaegher, Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Science*, *14*, 441-447.

Di Paolo, E. (2009). Extended life. *Topoi*, 28, 9-21.

Di Paolo, E., Rohde, M., & De Jaegher, H. (2010). Horizons for the enactive mind: Values, social interaction, and play. In J. Stewart, O. Gapenne, & E. Di Paolo (Eds.), *Enaction: Towards a new paradigm in cognitive science*. Cambridge, MA: MIT Press.

Horwich, P. (1998). *Meaning*. New York: Oxford University Press.

Kripke, S. (1982). *Wittgenstein on rules and private language*. Malden, MA: Blackwell.

Lo Presti, P. (2013). Social ontology and social cognition. *Abstracta*, *7*(1), 5-17.

Maturana, H., & Varela, F. (1980). *Autopoiesis and cognition: The realization of the living*. Dodrecht: Kluwer Academic Publishers.

Millikan, R. (1990). Truth rules, hoverflies, and the Kripke-Wittgenstein paradox. *The Philosophical Review*, *99*(3), 323-353.

Peacocke, C. (2012). Understanding and rule-following. In A. Coliva (Ed.), *Mind, meaning, and knowledge: Themes from the philosophy of Crispin*. Oxford: Oxford University Press.

Pettit, P. (1990). The reality of rule-following. *Mind*, *99*(393), 1-21.

Pettit, P. (2002). *Rules, reasons, and norms*. New York: Oxford University Press.

Soames, S. (1998). Skepticism about meaning: Indeterminacy, normativity and the rule following paradox. *Canadian Journal of Philosophy*, *23*, 211-250.

Thompson, E. (2007). *Mind in life*. Cambridge, MA: Harvard University Press.

Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind*. Cambridge, MA: MIT Press.

Wittgenstein, L. (1953). *Philosophical investigations*. Translated by G. E. M. Anscombe. Prentice Hall, NJ: Basil Blackwell.

Wittgenstein, L. (1956). *Remarks on the foundations of mathematics*. In G. H. von Wright, R. Rhees, & G. E. M. Anscombe (Eds.). Translated by G. E. M. Anscombe. Oxford: Basil Blackwell.

Wright, C. (2001). *Rails to infinity: Essays on themes from Wittgenstein's philosophical investigations*. Cambridge, MA: Harvard University Press.

Department of Philosophy and Cognitive Science,
Lund University, LUX,
Helgonavägen 3D,
223 62 Lund,
Sweden
Tel. +46 (0)705 95 25 93
E-mail: Patrizio.Lo_Presti@fil.lu.se

Paper 4

An ecological approach to normativity

Abstract

It is argued that normativity is an embodied and situated skill that resists explanation in terms of rule-following. Norms are dynamic and negotiable, and are understood in practice by engaging with others. Rules are a subclass of norms and have pragmatic functions, e.g. to impose norms and elucidate implicit normativity. The propositional articulation of norms is secondary to normativity. Norms can be explained within the framework of ecological psychology as a particular kind of affordance that enables actions to be directly understood as correct. This view entails that the niche of human beings is inherently normative. Finally, the ecological account of normativity is used to elucidate the notion of rule-following.

Keywords

Normativity, know-how, embodiment, ecological psychology, affordances, rule-following

1 Rules and knowing what to do

Normative behavior is not always rule-following behavior. This means that not all norms are rules. In support of this it will be argued that one can understand norms governing a practice (a) without knowledge of rules that define the practice, and (b) without intending one's action to accord with rules that define the practice. Normativity is grounded in participation and engagement, as argued by Brinck (2014).

The meaning of 'norm' and 'normative,' as I use these terms here, will initially be "a very basic one" in Rietveld's (2008a, p. 974) sense that "is revealed when we distinguish … correct from incorrect … in the context of a specific situation." This basic sense will be elaborated later, when I discuss how people can learn not only how to behave correctly but also how, among correct behaviors, to behave as one ought or should. Moreover, as I develop an ecological approach to norms below, I will argue that norms are essentially social. Norms are primarily to be understood at the sociocultural level, to populations engaged together in social practices. In this section I argue for claims (a) and (b).

(a) The first point to be argued is that one can understand norms governing behaviors without knowledge of rules that define them as practices. Consider Searle's (2015; 2010, pp. 96-98) distinction between constitutive and regulative rules. A *constitutive* rule defines a type of action or activity and thus describes that there are certain things one does and does not do in the context of that action or activity (Brinck 2015, pp. 7-9; Searle 2007, p. 14). The activity would not have existed had it not been for the definition. For instance, chess is defined as the practice it is by the rules constitutive of the game, such as: *X* (a piece of wood) counts as *Y* (a pawn) in *C* (the context of chess) by the set of moves one is allowed to make with *X*(*Y*) in *C* according to the definition of the game (Searle 2006). A *regulative* rule, by contrast, does not define. It regulates actions and activities that depend on means to end-relations in the context of an existing practice, whether or not that practice is naturally occurring or was created by the formulation of a constitutive definition (Brinck 2014, p. 738). Regulative rules standardize an activity relative to contingent properties that it acquires in contexts of interaction. To illustrate the distinction, traffic rules define driving in a community. But much behavior in the context of driving is not determined by the constitutive rules. For instance, one may listen to loud music while riding round and round a roundabout with the car windows down. Doing this is not to violate the constitutive rules. It can be regulated against, however, and constitute violations of the regulative rules that normalize driving behaviors.

I want to argue that insofar we are concerned with normative behavior the notion that it is an instance of following a constitutive rule is too restrictive (Brinck 2015). Specifically, I argue that normative behavior does not require inferentially articulated propositional thought. Norms regulate interaction between and among agents and groups of agents; they are implicit in sociocultural practices. Inferential articulation of constitutive rules (definitions) presupposes engagement. Definitional rules are in most cases *pragmatic elaborations* from basic, pre-conceptual, embodied and situated interaction in existing practices (Brandom 2008, pp. 26, 86). Constitutive rules play a limited role for normative behavior, as definitions of what one does, e.g., in the context of games or the grammar of a language. One may understand how to play the game or how to use language correctly without knowing about, or being able to state, the constitutive rules. These rules, therefore, play a limited role in accounting for an understanding of norms. To understand norms is to be able to *do* something that one may or may not already have the concept of (Brandom 2000, p. 81). In line with

Brandom's view, a capacity to articulate the rules allowing one to state *that* something is correct is not inherent in the ability to conform to norms. Rules function to make correctness conceptually explicit and presuppose an implicit normativity of the relevant practice that the rule is about.

To illustrate, consider a norm for standing distance in conversation. Describing an interlocutor's understanding of this norm in terms of rules we would say that she understands what distance is correct if and only if she has an inferentially articulated concept of 'conversation' (i.e. a grasp of definitional rules telling one what to do). This seems obviously false. Someone, such as a child, who has engaged in a practice many times, can learn to behave correctly without possessing the concept defining the practice. The child might do something in accord with norms without knowing that this is the case but which we, armed with inferentially articulated propositional contents, *call* "obeying" or "going against" the definitional rules (Wittgenstein 1953, §201; Brandom 2008, p. 46). This may occur simply because the person is an experienced practitioner of the practice we define.

Let us consider a set of necessary conditions that may plausibly be thought to apply to all cases of normative behavior.[1] In order for an agent's S behavior B to count as normative, at least the following conditions must be met:

(i) S must be able to differentiate between correctness and incorrectness of B relative to whatever normative considerations are in play,

(ii) the norms in terms of which S understands herself and her behavior must exert some influence on S doing B, and

(iii) S must be able to evaluate the legitimacy of the norms in terms of which she understands B.

On the conception of norms as implicit in social practices used here, conditions (i) and (ii) are satisfied. One must be able to differentiate what is from what is not normative and behave accordingly because of this understanding. The norms of the practice engaged must also exert an influence on what one does; S does B because she

[1] I would like to thank an anonymous reviewer at Adaptive Behavior for suggesting the following necessary conditions.

understands B as normative in her particular situation. Condition (iii), however, seems to require articulated knowledge of norms that apply, such that one may reflect on and evaluate their legitimacy. For this reason, condition (iii) requires what I below will call 'elucidative knowledge.' Elucidative knowledge, I will argue, is not necessary for normative behavior, in the relevant sense of knowing how to act correctly.

For illustration, consider Colin McGinn's (2011, p. 56; Searle 2011, p. 121; cf. 1983, pp. 150-52) example of a tennis player who learns through practice how to play. He then encounters a theoretically informed coach who tells him what it is that he is doing when he plays. According to McGinn, we can understand this as a case in which the player *already knows* everything he is told, but knows it in a different format: his knowledge, which is unarticulated and implicit, is *articulated* or *made explicit* by the coach (Brandom 2008, p. 110). The tennis player, like practitioners of other practices, might then learn how to say explicitly in propositional form what he already implicitly knew how to do. Let's say that what the player learns is, not new knowledge, but a way to articulate it; he gains *elucidative knowledge* (cf. Brandom 2000, p. 56).

The important point is that the acquired elucidative knowledge (stating that something is a norm) is secondary to and presupposes a pragmatic embodied unarticulated skill (knowing how to behave correctly). A tennis player may know how to differentiate between successful and unsuccessful strikes in such a way that this ability influences how he strikes the ball on a case-by-case basis without being in a position to evaluate the legitimacy of this differentiation in actual cases (something the coach can do). Analogously, someone may know how to differentiate norms in social practices in such a way that this knowledge influences what she does on a case-by-case basis without her also being able to reflect on the legitimacy of the norm (something that someone with elucidative knowledge can do).

Given this distinction between two kinds of knowledge (see also Hutto 2005 and Ryle 1949, pp. 16-20), the argument is that understanding and conforming to the norms of a practice requires one to know how to behave correctly but not to possess elucidative knowledge allowing one to say that something is correct (Heras-Escribano and de Pinedo 2015, p. 6). That is, it is not necessary that one's understanding of norms is elucidative. One can understand and learn how to behave correctly by way of participation in social practices without intending one's behavior to accord to the rules (if any) that define it as a practice. On this participatory construal of normative behavior one may also learn to behave correctly in the context of a naturally occurring practice,

i.e., one that is not previously defined, without being able to explicitate as a rule in propositional form what is normative in the context of the practice.

Nor is elucidative knowledge sufficient for understanding norms. The philosopher who can state everything that one must be able to do in order to play tennis might on that account still not know how to play tennis. He or she can have elucidative knowledge of what is meant by "net-roll" and speak lengthily about what one must be able to do to hit a net-roll, but the statement that this elucidative knowledge is sufficient for the ability to hit a net-roll is, it is plausible to assume, mistaken. As Heras-Escribano and de Pinedo put it, "one does not master the technique of driving or playing football if one just knows what [rules are] written in a book. One should know how to do it" (2015, p. 6). On analogy, one does not master the technique of behaving in accord with norms if one just knows what is said about what the norm is; one should know how to do what is normative. In short, the present claim is that normative behavior does not entail and is not entailed by elucidative knowledge. In the context of engaging in social practices one can learn how to behave correctly and become skilled at differentiating correctness from incorrectness on a case-by-case basis by way of participating in the relevant practice.

(b) The second point to argue is that one can behave in accord with norms governing a practice without intending one's action to accord with rules that define the practice. Knowledge of norms is embodied and situated. By this I mean that the skills for behaving correctly in the pursuit of a particular practice are based in, and realized through, bodily engagement (Dreyfus and Kelly 2007).

To exemplify, children and immigrants can learn norms for standing distance in conversation through intersubjective engagement—by physically positioning oneself as an embodied being in the presence of another and addressing or being addressed by him or her (Gallagher and Bower 2013, pp. 238-39). Being guided by others via negative and positive feedback allows the child and the immigrant to develop a non-conceptual 'feel' for correct standing distance. By engaging in conversations they with time gain participatory knowledge how to behave correctly. This does not require them to have a prior intention to learn the norm or to behave in accord with what a rule would express as correct. What is required for the acquisition of an understanding of norms is *participation*: making sense of the situation by physically putting oneself in it and taking part in it with others (Brinck 2014, pp. 742-49; Di Paolo and De Jaegher 2015, p. 1; Fantasia et al. 2014, p. 5). Processes of social participation, I will argue below, also

tend to produce change and progress in normative behavior, why it is often the case that there is no norm how to behave in the context of a particular practice that pre-determines what will be correct. For this reason it is often *required* to participate in a practice in order to learn how to behave correctly in the context of any particular situation. In the process agents acquire an embodied and situated participatory knowledge.

Through regular participation in social practice, someone may learn details of the 'normative fine structure' of a certain activity—e.g. whether to raise one's glass to the height of the second topmost button of one's shirt or to the height of one's chin when raising a toast. This fine-tuned understanding is something a practitioner can become acquainted with without being able to state, in the form of a general rule, what one must do to toast correctly, but through embodied engagement in the activity of toasting (McGann et al. 2013, p. 206).

The distinction between elucidative and participatory knowledge is between kinds of knowledge, between knowing that some behavior is correct and knowing how to behave correctly. The latter might be called a social skill: a kind of knowing "'stored', not as representations in the mind, but as dispositions to respond to the solicitations of [social] situations in the world" (Dreyfus 2002, p. 367). It is not a distinction between poles on a continuum, where the former represents 'higher-level' and the latter 'lower-level' knowledge. Participatory knowledge might be high-level in the sense that it constitutes an almost invariable capacity to discriminate correct from incorrect behavior. According to Rietveld (2012, p. 214) this is what it takes to be an expert. Elucidative knowledge can be quite low-level in the sense that it might amount to an unreliable capacity to conceptually distinguish correctness from incorrectness, where being unreliable signifies that this capacity tends to result in a wide range of false judgments about norms. Thus one might have intricate elucidative knowledge but no participatory knowledge, and vice versa (or one might have both) (Heras-Escribano and de Pinedo 2015, pp. 6-7). Having one is neither necessary nor sufficient for the possession of the other. Hence it seems that elucidative knowledge of the constitutive rules that define a given practice and the ability to propositionally explicitate rules, on the one hand, and embodied, situated participatory knowledge acquired via participation in the concrete context of a social practice, on the other, are *conceptually unrelated*. However, having either can *help* one to acquire the other (cf. Rietveld 2012, p. 208).

I have argued that understanding whether a given action or activity is correct or not does not imply being able to articulate the constitutive rules (definitions) of the practice. Having the concept of a certain practice is not necessary, nor is it sufficient, for understanding it (Moyal-Sharrok 2013, pp. 264-65). A situation in which a child or an immigrant learns how to position themselves correctly in conversation by way of participation is intelligible; and it is intelligible that someone who is given a rulebook stating how to engage a practice correctly may still fail to behave accordingly when engaging in the practice in particular situations (Brinck 2014, p. 740; 2015, p. 12). For example, someone may recount the rule that moving a pawn sideways in chess is wrong, because he has read the rulebook, and yet not know how to behave in a game of chess, because he has never played. Perhaps one ought not spit one's opponent in the face, for instance, but doing so is not to not play chess. The converse is also true. Knowing how to behave in the actual enactment of a practice is neither necessary nor sufficient for elucidative knowledge of its constitutive rules, i.e. for possession of the defining concept. For example, one may know how to behave in a game of chess (e.g., one should not spit on the board) without knowing the rules that define the game.

In the next section I will develop, in outline, an account of normativity as fundamentally dynamic.

2 The dynamicity claim

The claim that normativity is dynamic has a negative and positive part. First, it points to a problem with the picture of normative behavior as an instance of following constitutive rules. Secondly, it presents an alternative picture, on which rule-following is rather a special case of normative behavior. It will be argued that norms are not definitions of practices that, once laid down, determine future behavior. In line with Brinck (2014, p. 745), I will describe norms as established patterns of behavior in social interaction. Embodied and situated engagement has the capacity to change norms on a case-by-case basis in the concrete context of the social practices engaged. Norms are contextually action-dependent and negotiable. This is why an understanding of norms must also be understood as an embodied and situated practical sensitivity to the unfolding dynamics of the here-and-now contextual particularities of practices (Heft 1989, p. 18; Kiverstein 2012, p. 746). To shed light on these ideas, I begin by clarifying the central notions of contextual action-dependence and negotiability.

By saying that norms are contextually action-dependent I mean that what people do in a social interaction is constitutive of norms in the context of the specific situation. For instance, suppose 3ft is the correct distance for interlocutors to stand apart when they are in conversation in a given cultural context. I understand this as follows: people, having kept that distance in embodied interactions in the context of conversation, have developed an embodied and situated know-how to position themselves in the context of conversation (Rietveld 2008a, p. 977). If a listener stands further away, this might elicit the feeling that she is not really engaging in the conversation, that she is disinterested and should not be addressed. Standing closer than 3ft might be construed as threatening, and the speaker may respond with evasive behavior. Contextual social interactions might also modify people's perceptions of what is normative. For example, if, over time, people develop a habit of standing closer (perhaps they live in a noisy environment) or further away (perhaps they are used to speaking very loudly), and if this makes them feel comfortable with a more intimate or remote distance, then these contextual actions of closing or widening bodily standing distances in conversation can become perceived as normative: actions that previously prompted negative feedback (e.g., "You are standing too far away!" or "You are standing too close!") now do not, in the specific context.

The context-dependence of norms allows norms to change in micro-timescales (Klaasen et al. 2010). For instance, in a conversation in a crowded city, ebbs and flows in traffic might require speakers/listeners to move closer or further away, to pause in the middle of sentences, or to exaggerate lip-movements in order to be understood. By these changes being dynamic in micro-timescales I mean that they are immediate, occurring in social engagement over seconds or even shorter timespans (ibid, p. 59). Micro-timescale changes are, as Heras-Escribano et al. put it (2015, p. 25), not tied down to "a linguistic maxim or rules, because that general rule just expresses the general aspects […] not the distinctive and particular features that the agent deals with in specific situations." Norms are caught in a 'tidal wave' of contextual action. They are *particular dynamic relations* in the context of social practices. This means that norms are characterized by continuous change, activity, or progress in the context of a particular practice between individuals. Norms as particular dynamic relations are products of processes of embodied interaction in the context of social practices that tend to produce continuous change, activity, or progress in those interactions. What, from the point of view of someone who has elucidative knowledge *that*, *in general*, 3ft is the

correct standing distance in conversation in the context of a particular culture can change radically on the level of micro-timescales in ongoing social interactions (Heft 1989, pp. 18-19). Such change is not to be understood as a *violation* of norms. This provides evidence against the picture of normative behavior as an instance of following constitutive rules, as found in studies in developmental psychology (e.g., Rakoczy 2007; Rakoczy et al. 2008; Rakoczy et al. 2009). Rules, as definitions of correctness determining what to do, are too static and general to capture the progressive movement that causes norms to change several times on micro-timescales without anyone violating or misunderstanding them. The dynamic picture of norms supports claims (a) and (b) of the previous section: normative behavior does not necessarily involve elucidative knowledge of norms or an intention to make one's behavior accord with rules that explicitate or define norms.

The other side of dynamicity is negotiability (Brinck 2014, pp. 742-43, 745-47). Norms extended *over time* in a cultural context are constrained by the particularities of embodied and situated contextual action *at instants*. For example, while it might be true that it is a norm in Scandinavia to trust the authorities, the accuracy of an extrapolation from this macro-level description to generalized Scandinavian authority trust-norms is constrained by changes in the concrete context of authority-citizenry interactions. Thus the micro-level contextual action-dependence of norms *at a time* suggests that norms extending *over time* are negotiable. The negotiability of norms is a consequence, again, of embodied and situated social interaction. Change in norms at specific times tends to produce change in norms over time (Brinck 2014, p. 749). This reminds of the type of context-sensitivity of normative behavior that Rietveld (2008a) calls situated normativity.

That norms are negotiable, then, means that what from the point of view of someone who has elucidative knowledge that, in general, it is correct in the context of a particular culture to, for example, trust the authorities, might change over time as a function of the negotiation of norms at times. Rules, understood as generalizations over times attempting to explicitate norms for future behavior by extrapolation from present norms, cannot capture this dynamic process (Myin and Hutto 2015, pp. 61-62). As a result of the dynamicity in social interaction at temporal instants, the temporal extension of norms is underdetermined by present norms. Hence someone's statement that such-and-such behavior is normative is not only possibly falsified by what people do in context (context-dependence), but is also, as a statement of what will be correct, indeterminate

in truth-value because its accuracy depends on what people do at various instants over time (negotiability). The dynamicity claim is, in summary, this: norms for social practices, which may or may not also be defined by constitutive rules, can change several times both over seconds and over longer stretches of time. Because of the tendency of embodied social interaction to produce contextual agreement and disagreement in behavior, it tends to produce continuous change and progress among norms. Constitutive rules, e.g., those for chess or grammar, may be more static as definitions of practices, but norms for how to play or speak in particular situations in the context of practices defined are fundamentally dynamic. To illustrate, suppose that two friends are playing chess. One of them is a novice and the other is a professional player. To even the odds they agree to the latter not being allowed certain advanced tactics. They are still playing chess but they have negotiated norms for how to play in the particular context of this game.

Statements about norms answer to what people actually do, and only secondarily to what people say or believe one must do in order to behave correctly. Understood as definitions serving the pragmatic, explicitating function of making explicit norms that are implicit in embodied and situated action (Brandom 1994, pp. 21-22), rules-statements are at best approximately accurate extrapolations over macro-timescales covering present and future norms. This is because norms are relative to cultures and times.

Rule-expressions may also serve the declarative function of saying "X counts as Y in C" (Searle 2015, p. 510). Rule-expressions are used in this manner not to make implicit norms explicit, but to (attempt to) *make* some behavior correct; to define in what a given practice consists (Brinck 2015, passim; Searle 2010, pp. 97, 105-06).

On the dynamicity claim, declaring that norms apply and articulating them can be understood precisely as part of the dynamicity of norm progression. Norm-imposing declarative speech acts can be understood as attempts either to collapse or direct the progression of normative behavior. For example, by saying that something is to count as a correct move in the context of a game, one may, given others' acceptance of the declaration and subsequent conformity of action, succeed in making part of the game's normative structure static and nonnegotiable. If successful, norm-imposing rule-expressions may thus impose normative structure on behavior (Searle 1995, pp. 100-01). Certain behaviors become defined as obligatory or prohibited within the context of the rule. Over time this structure may transform the way people perceive the conditions

for social interaction; where before there was an undefined, inarticulate set of norms, now there is a well-defined system—an institution—to which people may or may not adapt.

What this shows is that the possibility, within a community, of elaborating a language expressively powerful enough to speak about norms introduces an even more radical dynamicity. I argue that both explicitating and norm-imposing functions of rule-expressions are special cases of a general norm-dynamics. After all, if a certain speech act can change normative behavior in the manner suggested, then extensive changes to norms can be sudden and widespread indeed. For example, if someone, regarded as an authority on these matters, says that from now on it is to be counted as a norm for bankers to carry a necktie at work, and if people accept this declaration and adapt their behavior to it, then the normative consequences can be sudden: from the next morning most or all bankers put on a necktie when they go to work. If they do not, their costumers and colleagues may think unfavorably of them. What this illustrates is that also attempted imposition of norms by declarative rule-expressions is a dynamic force; it tends to produce change in normative behavior. Because of this we can say that although definitional rules might impose a static structure of norms, this imposition is itself part of a process characterized by a general norm-dynamics.

Nevertheless, the extent to which definitional rules can change normative behavior may be questioned, as follows. What is really making a difference to norms in these cases is not people's understanding of the rules stating that something is or is to be counted as normative, considered in abstraction from the relevant social practice. What makes a difference to the norms is what people do, in this case as a consequence of the rule being expressed. Although the pragmatic purpose of rule-expressions may be to impose norms, the *success conditions*—what is required for the pragmatic purpose to be achieved—is that people behave in ways that the statement represents as correct because it is so represented (Searle 2010, pp. 12-13). So, for instance, for the declaration that it is a norm for bankers to carry neckties at work to have force it is not sufficient that people understand the declaration and can recount it. It is necessary that people do something as a consequence of the declaration and starts to actually treat it as a norm (they adapt, and if they do not they are sanctioned in some way). In this manner, the extent to which the imposition of norms by means of rule-expressions can, by themselves, change normative behavior may be questioned. If rule-expressions change

normative behavior then this is precisely a *special case* of producing change; it is a special case of continuous change and progress.

This special case-argument is important because it shows that we may incorporate rule-expressions in the dynamicity claim, as follows. They are speech acts that, as definitions of actions and practices, tend to produce new activities or changes in already existing ones, conditional on people adopting and adapting to them by acting. This is precisely a dynamic process.

Let us consider a concrete example of norm-dynamics, involving progress both implicit in behavior and explicit in rule-expressions. Suppose that, in a certain community, littering in public spaces is a norm. This, we may suppose, is a consequence of people in the community having established a social practice of littering. There is no explicit rule stating that littering is correct. Introduce into this community a certain merchandise, e.g., a kind of healthy chocolate bar, which becomes so popular that people eat it in vast amounts. As a consequence the wrapper of this chocolate starts to flood the streets. People then become attentive to how littered their environment is. They find it appalling. Because of this they start to throw their chocolate wrappers in suitably placed litterbins (indeed, they start to throw other waste in these bins as well). Time passes and the environment of this community gets cleaner and cleaner. After a year or so, littering behavior is met with social sanctions, e.g., scolding and disparaging glances. A littering-norm has implicitly been negotiated, because of contextual factors, and changed into a no littering-norm. Non-adapted behavior is sanctioned. Now consider a variant of this story. When the streets become more and more cluttered with chocolate bar wrappers the sanitation-authorities issue an edict stating that littering in the streets will be regarded as wrong (perhaps punishable). That is, littering becomes defined as incorrect. As a consequence the littering-norm may or may not change, depending on what people do; on whether they adapt to the edict. The edict has the potential to produce a sudden (or slow) change in the behavior of the community (this is indeed the whole point with the edict; to produce change). If people do what is declared as correct because it is so declared, then the edict has successfully produced a change in norms in the context of the practice of littering in public spaces. Hence we may understand the edict, the rule-expression, as part of an ongoing progression of norms by virtue of its tendency to produce change in normative behavior. Now suppose that, over time, people forget about the edict and resume their littering behavior. As a consequence the streets become cluttered once again and people no

longer perceive a no littering-norm as applying in their community. The norm is being renegotiated. What this story is meant to illustrate is that the imposition of norms by means rule-expressions is itself part of an ongoing dynamic process. It also illustrates that such imposition is not necessary for behavior to be normative or for normative behavior to change. In the first version of the story the littering-norm changed because of implicit progression. Nor is imposition sufficient, because people may not do what is declared as correct, and they may negotiate imposed norms over time.

To recapitulate, I have argued that norms are context-dependent and negotiable—i.e. in constant change both *hic et nunc* as a function of interactions taking place within the practices with which they are associated, and over time as a function of the patterns of social interaction that constitute normative behavior. Rules, it has been argued, have two functions: they can be approximate explicitations whose pragmatic function is to let us *say* that something *is* normative, as Brandom puts it. When serving this function, rule-expression may be more or less accurate, depending on whether the explicitation accurately represents norms implicit in social practice. To successfully use rule-expressions with this function an agent must have elucidative knowledge of the norms, otherwise she cannot make them explicit. Rules can also serve the imposing role of saying that something *is to be counted as* correct, as Searle puts it. When serving this function, rule-expressions can be more or less successful, depending on whether what is declared as correct produces a corresponding change of behavior in the context of the relevant practice. To successfully use rule-expression with this function an agent must both be able to define a practice as one that can be engaged only under certain conditions as well as make others adapt. I have argued that the very point of rule-expressions, in both their pragmatic functions, are intelligible as part of a dynamic process of implicit and explicit change, imposition, abandonment and negotiation of norms within the concrete context of embodied participation in the social practices with which they are associated.

According to the dynamicity claim, rule-following behavior is a special case of normative behavior. In line with a broadly Wittgensteinian approach, to behave correctly is not to follow rules as "rails … laid to infinity" (Wittgenstein 1953, §218), or even as shorter stretches of rails that we have not yet walked, rigidly determining how to go on. On that picture an individual can understand a norm only if she has elucidative knowledge that something is, or that it will be, correct. On the present view, in contrast, norms are understood as the rails we continuously lay down, not as already lain down,

and rule-expressions are treated as attempts either to explicitate our railroading activities or to direct them. In both cases the accuracy and success of rule-expressions is determined only in the concrete context of laying the next rail: it is determined in embodied action in the context of social practices.

The dynamicity claim clarifies the sense in which norms and normative behavior are dynamic. In the next section this claim is elaborated within the framework of ecological psychology, where norms can be understood to be a special class of affordances.

3 Norms are social affordances

Ecological psychology examines the ways in which perceptually guided behavior depends on properties of the environment in which it occurs. Agents in their environment constitute *eco*-logical systems, as contrasted with purely *physio*- or *psycho*-logical systems, in the sense that perception and action is constrained and enabled by the way agents' embodiment and the environment's structure together specify possibilities for perceptuomotor activity (Gibson 1979, p. 129). Thus in order to understand perceptuomotor behavior it is to the dynamics of embodied social interaction that we must turn—not to the external physics of a world, abstracted from agents inhabiting it, nor to the internal psychology of the agents, abstracted from the world they inhabit, including other agents.

A central explanatory concept in ecological psychology here is that of an *affordance*. James J. Gibson introduces it (ibid, p. 127) thus:

> The *affordances* of the environment are what it *offers* the animal, what it *provides* or *furnishes*, either for good or ill. The verb *to afford* is found in the dictionary, but the noun *affordance* is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment.

There has been a good deal of discussion about how, precisely, we are to understand the notion of an affordance.[2] In what follows I will use 'affordance' in the *relational* sense. I mean by it, as in the closing lines of the excerpt from Gibson, the complementarity of a group of agents and the properties of their environment. I will argue that affordances can also be understood to provide or furnish *normative* behavior.

[2] See, e.g., Zipoli Caiani (2013), Stoffregen (2004), Scarantino (2003), Chemero (2003), Jones (2003), Sanders (1997), Greeno (1994), and Turvey (1992).

Notice that I say that affordances relate to group of agents and their environment. I do so because I will focus on affordances that do not imply any particular *individual* and its environment. I thus take the position of Rietveld and Kiverstein, who think of affordances as embedded in sociocultural practices (2014, p. 326). This understanding of the affordance concept also agrees in outline with Stoffregen's (2004), in that it advocates a broad understanding; one involving not what an individual can do but what individuals do together in their shared environment.

According to Gibson, the set of affordances of a kind of animal in relation to its environment constitute that kind of animal's *niche* (1979, p. 128). The niche of a group of agents is not their physical environment, or *where* they live. Nor is it the phenomenal environment as privately, subjectively experienced. It is *how* they live—their shared "ways of life" (cf. Wittgenstein 1953, §241). Gibson (ibid, p. 128) writes:

> The natural environment offers many ways of life, and different animals have different ways of life. The niche implies a kind of animal, and the animal implies a kind of niche. Note the complementarity of the two.

Gibson's project of ecological psychology thus aims to understand perceptuomotor activity from the point of view of what the niche of a kind of animal affords its inhabitants, as a function of how agents are embodied and the structure of the environment (Withagen and van Wermeskerken 2010).

As Rietveld and Kiverstein note, Gibson sometimes speaks of affordances in social forms of life as "the whole 'spectrum of social significance'" (2014, p. 327; quoting Gibson 1979, p. 128). For instance, Gibson (ibid, p. 42) asserts: "what the other animal affords the observer is not only behavior but also social interaction." What agents do in social interaction is create a niche, a functional way of life from within the concrete context of which they directly perceive *how to act*, because on this theory perception is of affordances; perception is *for action*. Shared niches also take on social significance that allows agents to directly understand how to engage in social interaction (Gallagher and Bower 2013, pp. 237-38). We can thus speak of participatory know-how in the context of a niche as guiding engagements in a way of life shared with others (Heft 1989, p. 13), and of perception being "*for inter-action*" (Gallagher 2011, p. 68).

I suggest that Gibson's notion of a niche should be expanded so that it refers not just to functional ways of life, to how organisms live, but also to normative ways of life, to

how to live correctly. In this I follow Rietveld and Kiverstein (2014) in arguing that affordances provide a normative dimension associated with social practices. For example, norms associated with linguistic practices do not depend on any one agent but rather on the linguistic practices of a population of agents. Likewise, norms associated with standing distance in conversation depend on a population's engagement in that practice. This holds for practices in general. What is afforded as correct in the context of a game of chess, in driving, in doing mathematics, etc., does not depend on what is correct according to an *I* but on the population, the *we*. A game of chess will not afford moving a pawn sideways in my niche only because I decide to regard it as correct. Also, this means that if, e.g., a particular agent dies or disengages a social practice with certain affordances associated with it, these affordances do not disappear. The reason for this is that insofar the relevant affordances are associated with the way of life of a population the affordances continue to apply, despite some members' disappearance. Because of this, the notion of affordance as here understood has a social aspect; namely, that of how groups of people live together in a shared niche.

According to Heft, higher order ecological structures arise from the collective activities of individuals in conjunction with affordances. Heft (2013a, p. 165) writes:

> When individuals come together for a common purpose, they must agree, usually tacitly, to adhere to a set of socially normative practices so that their shared and individual intentions can be realized.

These practices "create possibilities for individual actions that only exist as a result of what these collective structures afford" (ibid). I propose that norms can be modeled as dynamic affordances, understood as relations among members of social groups and their shared environment. Embodied and situated understanding of normativity is the capacity of agents in the concrete context of a niche to adapt to and directly understand, in social interaction, what behavior is normative.

To see how this works, let us take the normative structure of a niche to be such that embodied social interactions at particular times establish and change what is correct in its context over time. Norms are then describable, not as part of the environment considered in abstraction from what agents do in social interaction, nor as part of what agents think or believe about what to do in abstraction from social interaction, but only as relational properties of their niche, regarded as an eco-dynamic whole (Heft 1989,

pp. 14-15; 2013b, p. 17). The eco-dynamic whole affords behavior that is socially normative. It thus seems to be correct that, as Heras-Escribano and de Pinedo claim, "the normativity of action must be *socially* mediated" (2015, p. 8; emphasis added). Normative predicates, correctness and incorrectness, are attributed to social interaction in the context of an embedding niche. The reason why I take this position is, as is Heras-Escribano and de Pinedo's (ibid, p. 16), essentially Wittgensteinian; namely, otherwise an individual who, in privately taking some behavior to be a norm, would thereby make it a norm, and thus it would not be possible to distinguish what is the norm from what one takes the norm to be (Wittgenstein 1953, §202). Therefore, we may speak of norms as social affordances, emphasizing that these affordances do not relate to any isolated individual in an embedding niche, but to a social group and the niche its members share. As Costall (1995, pp. 477-78) argues, social affordances are no less real on the ground that they are social. Insofar animals live together in an environment shaped by their social interactions into common niche, the affordances of the niche consist in relation between its inhabitants, presenting possibilities for social interaction (Gibson 1979, p. 42).

This social dimension is constitutive of everyday engagements. Thus, although not all affordances are normative, much of a niche of affordances is socially constituted and normative. What agents encounter in the concrete context of a niche is a dynamic system of norms rather than mere behavior to which norms must then be added by interpretation or inference. Situating the dynamicity claim in the wider context of ecological psychology allows us to describe norms in specific terms: norms are relational, action-dependent affordances in the concrete context of a social group's niche (Rietveld and Kiverstein 2014). People can learn how to respond directly to these affordances with respect to the relevant niche.

The present account of norms and normative behavior offers tools for explaining behavior that might be inexplicable if niches are understood only descriptively in functional terms of *how* people live, as Gibson did. For instance, sudden and gradual changes in behavior in a population can be attributed to the dynamics of social, embodied, contextual interactions that negotiate norms, or to authoritarian, declarative norm-impositions. Normative affordances are not directly perceivable, in contrast to ordinary affordances that present what one *can* do, such as those of seeing a stone as a paperweight or as a projectile. Nevertheless, normative affordances can be inferred as causes of changes in behavior of a population. For example, if we observe a population,

e.g., abandoning a practice of littering in public spaces, but we do not observe any change in the functional properties of their environment that might explain this change, then we may infer instead a dynamic process producing change in their normative behavior, or perhaps a rule has been expressed that people adapt to and which produce the observed change. As Heft puts it in the excerpt above, a higher order ecological structure of normative affordances is established as people tacitly agree in patterns of behavior in social practice. The norms may not be directly observed, but changes in behavior not explained by observed changes in the functions of a natural environment may be attributed to tacit norm-dynamics that people adapt to.

Let us summarize the argument thus far. The central notion is that norms are dynamic. By this I mean that norms are established and negotiated in embodied social interaction over time as well as at times such that certain behaviors can be directly understood as correct or incorrect relative to whether they deviate from, or conform to, the established pattern in the context of a niche. Deviation may elicit sanctions and other negative feedback. People learn how to behave correctly and how to avoid behaving incorrectly by participating in social practices with others in the concrete context of a shared niche. That is, understanding norms is a direct and participatory know-how (Brinck 2014, p. 745; Heft 2003, pp. 172-73). It does not require elucidative knowledge of rules stating that some behavior is, or is to be counted as, correct. This suggests an ecological interpretation of the dynamicity claim. On this interpretation norms are a special kind of affordance. What defines this kind affordance is that agents are presented with normative behavior and not only possible behavior. These affordances are related to which among possible behaviors are normative in the context of social practices relative to a niche. This is not to say that all affordances are normative (Chemero 2009, p. 218). It is to say that norms can be considered to be social affordances (Heras-Escribano and de Pinedo 2015, p. 23; Heft 2013a).

Related to the above, I want to emphasize two important issues.[3] First, it should be noted that norms are *relative to niches*. What is correct in one niche can be incorrect in another, and yet an individual can simultaneously inhabit both. For example, someone might be a board member of the Department of philosophy where he works, but also the lead singer of a rock band, and father of three. Norms applying in these contexts can be different. On the account developed here, this is as expected. What people do together

[3] I want to thank an anonymous reviewer at Adaptive Behavior for bringing the following two issues to my attention.

by participating in different social contexts is to establish and negotiate norms relative to those specific contexts. Knowing how to behave in accord with norms can thus be understood as a capacity to deal with correctness as it is shows up in particular situations in different social contexts. To acquire this kind of knowledge one must participate in the different social contexts. In this way one acquires a practical discriminative capacity to differentiate between and participate in the shaping of norms of the niche, and also to maneuver between different niches and to participate in the different patterns that establish their different norms. This is exhibited in the how, e.g., our board member philosopher who is also a rock singer and father of three differentiates the inappropriateness associated with howling at his colleagues from the appropriateness associated with howling into the microphone, and in how he differentiates the inappropriateness associated with issuing a curfew on his band-mates from the appropriateness of doing so on his children. According to the argument I have presented, a consequence of this 'overlapping' of niches and their members is that if our Department board member, rock singer, and father of three disassociates from, e.g., his rock band for some time, then this may result in deterioration of his understanding of norms associated with its practices, as exhibited when he re-engages. This is so because the others can, during his time away, negotiate the norms that earlier applied. He might know how to say everything, in terms of explicitating rules, that was regarded as correct when he disengaged from the band, and yet this may not be sufficient for him to be able to understand the norms when he returns. But it is not necessary for him to learn to state the new explicitating rules, because he can, by re-engaging with the band and participating in its activities, learn how to behave correctly. This illustrates once again that norms are dynamic; they can change silently and unmoved by what someone with elucidative knowledge, who does not engage in the practice, might state as "'obeying the rule' or 'going against it' in actual cases" (Wittgenstein 1953, §201).

Secondly, it should be emphasized that, in certain contexts, the claim that behavior is correct may mean that it is *successful*, given means-ends considerations in the particular circumstances, while in other contexts it might mean that it is something one *ought* to do. Among possible behaviors in the context of a particular niche there can be many correct behaviors such that a participant of a particular niche may have to be able to discriminate not just what is successful but what she ought to do among the many things that are correct to do. This raises the question how, e.g., a parent can single out among the many correct, in the sense of successful, ways of seeing to it that her children do not

use drugs, one way that is correct in the sense that it ought to be opted for. On the account here developed, this is the question how a niche can afford members normative and not only successful behavior in particular situations. To single out normative behavior requires subtler discriminative know-how (Rietveld 2008b, p. 344). One must single out correct from possible and then normative from correct behavior. This can be handled by noticing that the required subtlety can be attained by way of participation in a niche of social practices. It might involve extensive trial-and-error, as when a parent tries different methods to make a child stop using drugs and finally, hopefully, finds one that is just right (cf. Rietveld 2008a/b). Brochures and directives (stating that to do this-or-that has been successful in other particular circumstances) might help. But it is, on the approach here elaborated, in the concrete context of the niche of this particular family, in how parent and child negotiate behavior in their circumstances, that one procedure may be chiseled out as normative. For example, it might be correct, in the sense of successful, to lock the child up for detoxification. But this might not, in this family, be regarded as correct in the sense of how they ought to cope (it might also not be successful for this family). This subtlety in normative behavior provides further support for the dynamicity claim and the ecological approach to norms. What the parent and child in our example do is to negotiate which among possible behaviors in the context of their niche are correct means given their ends, and which among these they ought to pursue. Ordinarily, they would not first define what they do and then state what they could do, in order to define how to do it correctly, in order finally to formulate a norm prescribing one method, from which they infer what they ought to do. Importantly, even if this is how they proceed they will of course have to do something in order for the norm to be validated. For them to learn what they ought to do they must try to actually do it. In this it it is neither necessary nor sufficient to have the relevant corresponding elucidative knowledge. The subtle discriminative capacity is, in ordinary circumstances, a practical know-how acquired by means of extensive participation in the relevant practice.

In the next and final section I shall ask whether the concept of normativity elaborated here sheds new light on the notion of rule-following.

4 Blind rule-following and bedrock

I end by setting the eco-dynamic approach to the task of illuminating certain long-lived mysteries about rule-following behavior. The first mystery concerns the sense in which normative behavior is rule-following. The second concerns in what sense normative behavior can be said to be "blind." Finally, I give an eco-dynamic interpretation of the notion of "bedrock," supposedly non-elucidative, normative behavior. By going through these issues I also hope to position the eco-dynamic approach in the context of certain contemporary debates on normativity and what I have called elucidative knowledge.

Wittgenstein's notion of "blind" rule-following (1953, §219) has attracted a good deal of attention. This notion enters Wittgenstein's investigations in an argument that no interpretation of a practice as exemplifying a rule can determine how to go on because every interpretation can be made to both accord and conflict with the rule (§§198, 201). The reason for this is that an interpretation can itself be interpreted as being both in accord and conflict with the rule given an alternative interpretation. Knowing how to go on cannot be determined by a rule under an interpretation without leading to a regress (§199). For example, suppose you are asked why you keep a certain standing distance in conversation. You might want to answer by referring to a rule, e.g., "In our community this distance is correct." You are referring to an interpretation of the practice according to which it is governed by a certain rule. But the practice of keeping the relevant distance can be justified according to an alternative interpretation, e.g., that it is correct to keep a 3ft distance unless it is Tuesday and you are standing on a street with a name beginning with the letter "B." As it turns out, it is Tuesday and you are on Baker's Street. Nothing seems to support your interpretation in favor of the alternative (and a host of alternatives). For reasons like this, Wittgenstein says that, in the end, when we run out of justifications for our acting one way rather than another, we must acknowledge that we follow the rule "blindly."

I argue for a construal of bedrock as *social*, as a niche of social practices affording normative behavior such that no *individual* can be said to behave in accord with norms if the behavior is not situated in the context of a social practice. Also, one *need* not justify, in a sense requiring elucidative knowledge, one's behavior insofar it is part of a social practice.

To appreciate the argument we must first understand what, precisely, is meant by rule-following being blind. To this end, it is instructive to consider what Wittgenstein actually says. He says that when we cannot give a further justification of why we do what we do we must say: "'This is simply what I do'" (§217), "I obey the rule *blindly*"

(§219). So "blind" in the context of following a rule means: without being able to further justify, or to state a reason for, one's behavior. And yet, Wittgenstein says, what we do is not haphazard, because we act responsibly (§222). Even if we cannot explicitate a further justification for our behavior, others, with whom we engage in the social practice in question, can hold us responsible for it. People, with whom we share a niche of social practices, techniques, and customs how to act, can accept what we do as correct or incorrect in the context of the relevant niche (§§ 199, 241; cf. Gibson 1979, p. 128). Others can help us elucidate the norms of the niche when we fail. Because of this people may behave correctly without being able to state that it is so; people can rely on each other to hold each other responsible. We help each other learn normative behavior by engaging with each other in social practices in particular situations in the context of our niche. In so doing we relieve each other from having to have elucidative knowledge of norms.

So when I say that normative behavior is not rule-following behavior I mean by this that there is what we may "*call* 'obeying the rule' or 'going against it' in actual cases" (§201; emphasis added), but which is, fundamentally, negotiated in actual cases of participation in a social practice.

Still, people may speak or think about what *would* be correct at a given instant if what is *now* correct were put 'on freeze.' Thus they might say or think something like: "Three feet is the correct standing distance in conversation, *ceteris paribus*…"—e.g., as long as you are a member of this particular community, are not in a heavily trafficked area, no participant suffers a hearing impairment, you are not whispering a secret, it is not Tuesday, and you are not on a street whose name begins with the letter "B", and… and so forth. *Ceteris paribus*-clauses like these are in principle indefinitely expandable because of the context-dependence and negotiability of norms discussed earlier. There is no room in this story for the propositional content of generalized rules, stating that something is correct, to capture norms as static over time, i.e. as applying to all particular situations. To behave correctly is not to follow a rule, as traditionally understood, but this is what people who can reflect on norms can call it. Of course, however, some practices, like speaking intelligibly in accord with rules constitutive of grammar, or to play chess according to the rules that define the game would still, in a sense, be instances of following rules. But also in these cases it is important to recognize two things.

First, one can correctly engage in these 'static' practices without knowing how to explicitate norms as propositional contents of rule-expressions, or how to justify one's behavior by explicitly recounting (an interpretation of) a rule stating that what one does is correct. For example, the tennis-player, the child who learns to play chess, the immigrant who learns how to position herself in conversation in her new circumstances, may all learn and know how to behave correctly not because they have elucidative knowledge of a rule but because they have participated in the relevant practices sufficiently to know how to behave.

Secondly, as was obvious earlier in the discussion of Searle's distinction between constitutive and regulative rules, there are many things that might be correct or incorrect within the context of playing or speaking according to constitutive rules. Not to spit one's opponent in the face, to shake one's opponents hand after the game, and so on, are things that within the context of playing according to the rules may or may not be regulated. Hence, even if one knows the constitutive rules stating that one must do this or that in the context of a social practice *as defined*, knowing this definition is neither necessary nor sufficient for understanding how to behave within the context of the rules. Thus I can play chess correctly, although I make frowns and spit on the board, because I successfully follow the rules that define chess. Clearly, though, in this case and others similar to it, I might also be playing incorrectly because it is taken to be inappropriate, something I ought not, to frown and spit on the board.

I also claim that knowing how to behave correctly is not "blind," in Wittgenstein's sense, clarified above. At least, it should be carefully reconsidered just how blind normative behavior really is.

Norms, as affordances, are part of a niche one directly responds to in the context of social practices of a niche (cf. De Jaegher 2015, p. 124; 2009; Gallagher 2004). A niche affords normative behavior—not just what it is possible to do but how to do it correctly (cf. Gallagher 2011, p. 66; Gallagher et al. 2013). If one has participated enough in a niche to learn how to behave in accord with norms, one has the participatory, embodied, and situated skill to take other people to afford not only social interaction but also normative social interaction. Thus norms are part of what a skilled participant with subtle discriminative capacities, as discussed earlier, can learn to respond to directly without interrupting interpretations or inferences.

Hence niches, including affordances how to live correctly together with others, are steeped, as it were, in normativity (Heft 2007, p. 92). The notion that normative

behavior is "blind" rule-following behavior must not blind us from the fact that understanding how to behave correctly is a direct responsiveness to norms considered to be affordances. The ecological dynamicity claim helps us see this understanding of norms as a direct understanding. Perhaps this is what Wittgenstein meant by acting responsibly without further justification. If it is, then the notion that to behave correctly is to blindly follow a rule is somewhat misleading. This is so because, first, we might not be following rules when we behave correctly and, secondly, because of course it is trivially true that we do not literally see norms. Nevertheless, in a non-trivial sense, we do see norms; namely, in the sense that norms are directly presented to the experienced participant in a niche of social practices. They can be directly understood even if not literally seen.

The alleged blindness of rule-following foreshadows another of Wittgenstein's remarks, mentioned above: that of "bedrock" action (1953, §217). Here "bedrock" refers to the point in our cascade of justifications of our interpretation of a rule where we admit that we can give no further interpretation. We should therefore not think of rule-following behavior as behavior premised on an interpretation (§201) of a rule. Interpretations can be made both to accord and to conflict with the supposed rule, since the rule admits of many, indeed, an indefinite number of, possible interpretations, as discussed earlier (§ 198, 201; cf. Ryle 1949/2009, p. 18). Thus we reach a point where we hit "bedrock" and have to say that, when we follow the rule, we act blindly.

Much discussion of rule-following and skill centers on how to make sense of this notion of bedrock. For instance, it is asked whether something we do without justifications can really be *normative* (McDowell 1984, pp. 340-42; Tanney 1999, p. 59; Glüer and Wikforss 2009, p. 59) and whether intentionality at the level of bedrock activity is conceptual and propositionally articulate or rather a pre-conceptual and non-propositional embodied skill (Dreyfus 2012; 2013, pp. 18-19; Hutto 2012; McDowell 2013, p. 54).

From the point of view of the eco-dynamic approach to normativity, bedrock can be understood as the concrete niche that affords normative behavior. Bedrock activity is normative and conceptually articul*able*. That is, it is possible for articulate beings like us to make explicit—to have elucidative knowledge of—norms implicit in our social practices (Brandom 1994, pp. 267-69). This is one of the pragmatic points with norm-explicitating terms, like 'should' and 'ought.' They allow us to speak about norms implicit in our niche of social practices that might never be explicitated and that are not

already defined. From this point of view, an entire population could lack the expressive resources of normative terms, and thus be unable both to make bedrock-level norms conceptually explicit and to impose new ones, yet this would not deprive the behavior from being adapted to norms of their niche. In this case, however, the population would not know *that* something is normative and would not have elucidative knowledge of norms.

Hence norms, as dynamic affordances, are not necessarily unarticul*able*. The normativity of the niche is, indeed, realized in bedrock-level, ongoing, embodied social interactions, and it does not presuppose elucidative knowledge. The fact that a population inhabits a niche affording norms does not imply that any of its members knows that it does, or that, if they know it, they can articulate those norms. But also it does not imply that they could not do so. The fact that norms are realized at bedrock-level and can be learned by way of participation in the niche does not imply that the inhabitants of the niche cannot articulate and have elucidative knowledge of those norms. Therefore it is not a necessary condition on behavior being normative that any of the inhabitants of the niche conceptually articulate what it is that they are doing in a way that indicates whether they are behaving according norms of their niche. But also, it is not necessarily true that they cannot have elucidative knowledge of the norms. Thus, I argue, Dreyfus is right here. Bedrock-level activity can be normative yet non-conceptual.

Nevertheless, because of the essentially social nature of norms, an individual who has learnt from participation how to engage in bedrock-level activity in accordance with norms without elucidative knowledge may still participate in a conceptually articulated scheme of correctness. This is so because others, with whom she shares the niche, can stand proxy for her in articulating whether she behaves as is correct. Thus the individual who skillfully behaves in accord with norms implicit in practice is not required to have any concept of what she does, because what she does is part of a social practice within the context of which others may articulate the norms and, over time, help her to articulate them as well.[4] This individual behaves in a manner that a social group, a "we [who have the conceptual capacities] call 'obeying the rule' or 'going against it' in

[4] Brandom (1995, p. 902) criticizes McDowell precisely for failing to appreciate "the essential *social* articulation" of concepts and reasons. This failure would indeed lead to the individual being required to apply concepts when she acts. If the social articulation of norms is emphasized, then this requirement on the individual can be flouted and the individual still be said to behave as is normative; namely, because she is a participant to social practices together with others who can assist her in articulating the norms.

actual cases" (Wittgenstein 1953, §201), although her behavior does not involve her having or applying concepts or intending her behavior to accord with what we, with elucidative knowledge, would explicitate as a rule. The social articulation of norms implies that the subject who knows how to behave in accord with norms can do so without concepts. What she does can nonetheless be "conceptual already" in the sense that she stands on the shoulders of others who provide the conceptual scaffold within the context of which the individual's activities need not be pervaded by conceptuality (Dreyfus 2013, pp. 15-16).

For an individual considered in isolation from a niche of social practices, bedrock behavior is non-normative. Only when situated in a niche of social practices is it normative. But that it is normative does not mean that the individual is required to have elucidative knowledge, because what she does can be conceptualized as normative instead from the social point of view supported by her niche-companions. In this sense, bedrock-level behavior might be already conceptual (McDowell 2013, p. 43), but not from the individual point of view (e.g., a child or an immigrant, or anyone new to a practice who learns by participation). The conceptual realm involved in bedrock behavior under these circumstances does not drive a distancing wedge between the individual and the social practices in which she is engaged, because the individual is not required to have the conceptual capacities to articulate bedrock-level norms since these norms are essentially *social* already. The socialization of bedrock-level normativity allows for behavior to be immersed in an already socially articulated conceptual scheme without requiring conceptual capacities on the part of the individual. The sociality of implicit normativity closes the gap that conceptuality, according to Dreyfus (2007, p. 354), opens. Being held responsible by others one can learn how to behave and how to hold others responsible, and how to explicitate responsibility. Insofar behavior is normative it is part of a social practice, and insofar one participates in a niche of social practices one may learn to respond to the affordances how to behave in accord with norms through participating with others, whose companionship provides a conceptual background for one to be socialized into over time. This is a dynamic process. Participation tends to produce progress and change of norms over time. Thus also the newly initiated, if allowed to participate in the social practices of the niche, may herself, without having elucidative knowledge of this, negotiate and update implicit norms over time together with others while being socialized into the niche.

Where does this leave us on how to understand "bedrock" behavior and normativity? Bedrock is not to be understood as a mysterious rendezvous with an individual's "mere behavior" on one side and the individual's "mindedness" on the other (Heft 1989, pp. 5-6). Rather, bedrock is the niche of dynamically changing normative social practices that we create, negotiate, and adapt to in interaction. It is a socialized realm. Yet bedrock behavior may then be conceptual already courtesy the social scaffold provided by others whom do have the conceptual capacities that the individual is not required having. These others, e.g., caretakers, supervisors, or others experienced in how to behave in the context of a particular niche, can say what the newcomer can only do, and can teach the newcomer to speak as they do. As Wittgenstein puts it, they can call normative what someone already knows how to do.

I hope to hereby have cast new light on some of the more intractable mysteries about rule-following behavior and normative behavior. First, normative behavior is not always an instance of rule-following behavior because in order to behave in accord with norms one need not have elucidative knowledge of rules that explicitate or define a practice and then intend one's behavior to accord with a rule. Secondly, the sense in which bedrock normative behavior is "blind," i.e. without justification, should not mislead us. Although the individual might indeed be blind in the relevant sense of lacking concepts, she is, insofar we are concerned with norms, part of a niche of social practices within the context of which her companions may lend her "sight." This means that what she learns how to do might indeed already be conceptually articulated even if her understanding is non-conceptual. The blind may remain blind and still see, in the sense of knowing how to do everything that is normatively required of her, if she is allowed to participate in the light of others who can tutor her through the darkness. If social, also a whole population of conceptually blind individuals may negotiate a normative way for how to live together. In this sense also the blind may lead the blind in normative ways of life. A population whose social interactions enact or change behavioral patterns such that members must learn to participate, differentiate, and adapt in order to avoid social sanctions can be said to have a normative way of live. However, lacking elucidative knowledge, it would be very different from human normative life, as we know it with its institutions, definitions, and explicit regulations about what to do, and, most importantly, with its language that allow us to make these things explicit.

That affordances can be socialized has been appreciated already from the various perspectives of ecological psychology (Costall 1995; 2012), philosophy (McDowell

2007; Dreyfus 2005), phenomenology (Gallagher 2013; Kiverstein In Press), and the cognitive sciences (Hutto and Kirchhoff In Press; Kiverstein and Miller 2015). Systematic treatment of the way in which affordances are dynamic and normative is still in its infancy (Rietveld and Kiverstein 2014; Heras-Escribano and de Pinedo 2015). The present investigation is meant to contribute to this. In this contribution I hope to also have provided elucidative knowledge how to tackle certain intriguing questions about rule-following behavior.

Acknowledgements: I am indebted to Ingar Brinck and two reviewers at Adaptive Behavior for extensive criticisms that greatly helped to improve on an earlier version of this paper.

References

Brandom, R. (1994). *Making it explicit*. Cambridge, MA: Harvard University Press.

Brandom, R. (1995). Knowledge and the social articulation of the space of reasons. *Philosophy and Phenomenological Research, 55*(4), 895–908.

Brandom, R. (2000). *Articulating reasons*. Cambridge, MA: Harvard University Press.

Brandom, R. (2008). *Between saying and doing*. New York, NY: Oxford University Press.

Brinck, I. (2014). Developing an understanding of social norms and games: Emotional engagement, nonverbal agreement, and conversation. *Theory & Psychology, 24*(6), 737–754.

Brinck, I. (2015) Understanding social norms and constitutive rules: Perspectives from developmental psychology and philosophy. *Phenomenology and the Cognitive Sciences,* ONLINE 15 June 2015. DOI 10.1007/s11097-015-9426-y.

Chemero, A. (2003). An outline of a theory of affordances. *Ecological Psychology, 15*(2), 181–195.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: The MIT Press.

Costall, A. (1995). Socializing affordances. *Theory & Psychology, 5*(4), 467–481.

Costall, A. (2012). Canonical affordances in context. *Avant, 3*(2), 85–93.

De Jaegher, H. (2009). Social understanding through direct perception? Yes, by interacting. *Consciousness and Cognition, 18*, 535–542.

De Jaegher, H. (2015). How we affect each other. *Journal of Consciousness Studies, 22*(1–2), 112–132.

Di Paolo, E. A., & De Jaegher, H. (2015). Towards an embodied science of intersubjectivity: widening the scope of social understanding research. *Frontiers in Psychology, 6*, 1–2.

Dreyfus, H. L. (2002). Intelligence without representation – Merleau-Ponty's critique of mental representation. *Phenomenology and the Cognitive Sciences, 1*, 367–383.

Dreyfus, H. L. (2005). Overcoming the myth of the mental: How philosophers can profit from the phenomenology of everyday expertise. *Proceedings and Addresses of the American Philosophical Association, 79*(2), 47–65.

Dreyfus, H. L. (2007). The return of the myth of the mental. *Inquiry, 50*(4), 352–365.

Dreyfus, H. L. (2012). The mystery of the background *qua* background. In In Z. Radman (Ed.), *Knowing without thinking* (pp. 1–10). Basingstoke, UK: Palgrave Macmillan.

Dreyfus, H. L. (2013). The myth of the pervasiveness of the mental. In J. K. Shear (Ed.), *Mind, reaon, and being-in-the-world: The McDowell-Dreyfus debate* (pp. 15–40). New York, NY: Routledge.

Dreyfus, H. L., & Kelly, S. D. (2007). Heterophenomenology: Heavy-handed sleight-of-hand. *Phenomenology and the Cognitive Sciences, 6*, 45–55.

Fantasia, V., De Jaegher, H., & Fasulo, A. (2014). We can work it out: an enactive look at cooperation. *Frontiers in Psychology, 5*, 1–11.

Gallagher, S. (2004). Situational understanding: A Gurwitschian critique of theory of mind. In L. Embree (Ed.), *Gurwitch's relevance for cognitive science* (pp. 25–44). Dodrecht: Springer.

Gallagher, S. (2011). Interpreting embodied cognition. In W. Tschacher & C. Bergomi (Eds.), *The implications of embodiment: cognition and communication* (pp. 59-70). Exeter: Imprint Academic.

Gallagher, S. (2013). The socially extended mind. *Cognitive Systems Research, 25–26*, 4–12.

Gallagher, S., & Bower, M. (2013). Making enactivism even more embodied. *Avant, 5*(2), 232–247.

Gallagher, S., Hutto, D. D., Slaby, J., & Cole, J. (2013). The brain as part of an enactive system. *Behavioral and Brain Sciences, 36*(4), 421–422.

Gibson, J. J. (1979). *The ecological approach to visual perception*. New York, NY: Psychology Press.

Glüer, K., & Wikforss, Å. (2009). Against content normativity. *Mind, 118*(469), 31–70.

Greeno, J. G. (1994). Gibson's affordances. *Psychological Review, 101*(2), 336–342.

Heft, H. (1989). Affordances and the body: An intentional analysis of Gibson's ecological approach to visual perception. *Journal of the Theory of Social Behaviour, 19*(1), 1–30.

Heft, H. (2003). Affordances, dynamic experience, and the challenge of reification. *Ecological Psychology, 15*(2), 149–180.

Heft, H. (2007). The social constitution of perceiver-environment reciprocity. *Ecological Psychology, 19*(2), 85–105.

Heft, H. (2013a). An ecological approach to psychology. *Review of General Psychology, 17*(2), 162–167.

Heft, H. (2013b). Environment, cognition, and culture: reconsidering the cognitive map. *Journal of Environmental Psychology, 33*, 14–25.

Heras-Escribano, M., Noble, J., & de Pinedo, M. (2015). Enactivism, action and normativity: a Wittgensteinian analysis. *Adaptive Behavior, 23*(1), 20–33.

Heras-Escribano, M., & de Pinedo, M. (2015). Are affordances normative? *Phenomenology and the Cognitive Sciences,* ONLINE 15 June 2015. DOI 10.1007/s11097-015-9440-0.

Hutto, D. D. (2005). Knowing *what*? Radical versus conservative enactivism. *Phenomenology and the Cognitive Sciences, 4*, 389–405.

Hutto, D. D. (2012). Exposing the background: Deep and local. In Z. Radman (Ed.), *Knowing without thinking* (pp. 37–56). Basingstoke, UK: Palgrave Macmillan.

Hutto, D. D., & Kirchhoff, M. D. (In Press). Looking beyond the brain: Social neuroscience meets narrative practice. *Cognitive Systems Research*, ONLINE first: http://dx.doi.org/10.1016/j.cogsys.2015.07.001.

Jones, K. S. (2003). What is an affordance? *Ecological Psychology, 15*(2), 107–114.

Kiverstein, J. (2012). The meaning of embodiment. *Topics in Cognitive Science, 4*, 740–758.

Kiverstein, J. (In Press). Empathy and the responsiveness to social affordances. *Consciousness and Cognition,* ONLINE first: http://dx.doi.org/10.1016/j.concog.2015.05.002.

Kiverstein, J., & Miller, M. (2015). The embodied brain: towards a radical embodied cognitive neuroscience. *Frontiers in Human Neuroscience, 9*, 1–11.

Klaasen, P., Rietveld, E., & Topal, J. (2010). Inviting complementary perspectives on situated normativity in everyday life. *Phenomenology and the Cognitive Sciences, 9*, 53–73.

McDowell, J. (1984). Wittgenstein on following a rule. *Synthese, 58*, 325–363.

McDowell, J. (2007). What myth? *Inquiry, 50*(4), 338–351.

McDowell, J. (2013). The myth of the mind as detached. In J. K. Shear (Ed.), *Mind, reason, and being-in-the-world: The McDowell-Dreyfus debate* (pp. 41–58). New York, NY: Routledge.

McGann, M., De Jaegher, H., & Di Paolo, E. A. (2013). Enaction and psychology. *Review of General Psychology, 17*(2), 203–209.

McGinn, C. (2011). *Truth by analysis*. New York, NY: Oxford University Press.

Moyal-Sharrok, D. (2013). Wittgenstein's razor: the cutting edge of enactivism. *American Philosophical Quarterly, 50*(3), 263–279.

Myin, E., & Hutto, D. D. (2015). REC: just radical enough. *Studies in Logic, Grammar and Rhetoric, 41*(54), 61–71.

Rakoczy, H. (2007). Play, games, and the development of collective intentionality. In C. Kalish & M. Sabbagh (Eds.), *Conventionality in cognition: How children acquire representations in language, thought and action. New Directions in child and adolescent development, No. 115* (pp. 53–57). San Francisco: Jossey Bass.

Rakoczy, H., Warneken, H., & Tomasello, M. (2008). The sources of normativity: young children's awareness of the normative structure of games. *Developmental Psychology*, *44*(3), 875–881.

Rakoczy, H., Brosche, N., Warneken, F., & Tomasello, M. (2009). Young children's understanding of the context relativity of normative rules in conventional games. *British Journal of Developmental Psychology*, *27*(2), 445–456.

Rietveld, E. (2008a). Situated normativity: The normative aspect of embodied cognition in unreflective action. *Mind, 117*(468), 973–1001.

Rietveld, E. (2008b). The skillful body as a concernful system of possible actions. *Theory & Psychology, 18*(3), 341–363.

Rietveld, E. (2010). McDowell and Dreyfus on unreflective action. *Inquiry, 53*(2), 183–207.

Rietveld, E. (2012). Bodily intentionality and social affordances in context. In F. Paglieri (Ed.), *Consciousness in interaction: The role of the natural and social context in shaping consciousness* (pp. 207–226). Amsterdam: John Benjamins.

Rietveld, E., & Kiverstein, J. (2014). A rich landscape of affordances. *Ecological Psychology, 26*(4), 325–352.

Ryle, G. (1949/2009). *The concept of mind*. New York, NY: Routledge.

Sanders, J. T. (1997). An ontology of affordances. *Ecological Psychology, 9*(1), 97–112.

Scarantino, A. (2003). Affordances explained. *Philosophy of Science, 70*, 949–961.

Searle, J. (1983). *Intentionality*. Cambridge, UK: Cambridge University Press.

Searle, J. (1995). *The construction of social reality*. New York, NY: The Free Press.

Searle, J. (2006). Social ontology: Some basic principles. *Anthropological Theory, 6*(1), 12–29.

Searle, J. (2007). Social ontology: The problem and steps toward a solution. In S. L. Tsohatzidis (Ed.), *Intentional acts and institutional facts: Essays on John Searle's social ontology* (pp. 11–28). Dodrecht: Springer.

Searle, J. (2010). *Making the social world: The structure of human civilization*. New York, NY: Oxford University Press.

Searle, J. (2011). Wittgenstein and the background. *American Philosophical Quarterly, 48*(2), 119–128.

Searle, J. (2015). Status functions and institutional facts: reply to Hindriks and Guala. *Journal of Institutional Economics, 11*(3), 507–514.

Stoffregen, T. (2004). Breadth and limits to the affordance concept. *Ecological Psychology, 16*(1), 79–85.

Tanney, J. (1999). Normativity and judgement. *Proceedings of the Aristotelian Society, Supplementary Volumes, 73*, 45–61.

Turvey, M. T. (1992). Affordances and prospective control: An outline of the ontology. *Ecological Psychology, 4*(3), 147–187.

Withagen, R., and van Wermeskerken, M. (2010). The role of affordances in evolutionary process reconsidered. *Theory & Psychology, 20*(4), 489–510.

Wittgenstein, L. (1953). *Philosophical investigations*. Oxford, UK: Blackwell.

Zipoli Caiani, S. (2013). Extending the notion of affordance. *Phenomenology and the Cognitive Sciences*, *13*(2), 257–293.

Paper 5

Speaking about the normativity of meaning

Abstract: Contemporary debate on the nature of meaning centres on whether meaning is normative. While agreement is widespread that meaning implies correctness, analyses of correctness diverge on whether correctness is normative. Normativists argue that correctness implies obligations or permissions. Anti-normativists disagree and hold that correctness is a descriptive term. My proposal is that, fundamentally, meaning presupposes norms, but not in the generic normativist sense. For a vocabulary to be recognizable as part of a language it must be part of a practice of committing and entitling to ask for and provide reasons for what is said. To commit and entitle is not obliged or permitted though, but simply a presupposition for even speaking about obligations and permissions.

1. Introduction

According to a 'meaning platitude' (Hattiangadi 2006, p. 222), meaning implies correctness. Why so? Says Hattiangadi: we want to distinguish random noise from noise with meaning attached. The latter is assessable as correctly or incorrectly uttered, the former is not.[1] Lets assume that the meaning platitude is right.

What, now, constitutes correctness conditions for noise making so that we may distinguish noise with meaning from noise without? Correctness, it seems, is normative (Gibbard 2005, p. 338). Normativists argue that correctness conditions consist in obligations or permissions. Meaning is normative in that correctness conditions specify what one *ought* or *may* say.[2] Anti-normativists, in contrast, propose that correctness consists in

[1] Both advocates and opponents to the claim that meaning is normative now agree on the meaning platitude. See Hattiangadi (2006, p. 224; 2009), Boghossian (2008, p. 207; 2012, p. 37), Glüer and Wikforss (2009, p. 36), Whiting (2009, p. 537). The disagreement hinges on whether correctness should be nonnormatively or normatively construed—and, if the latter, what precisely a normative construal amounts to (Whiting 2013).

[2] This idea is familiar from Kripke's (1982) exposé of Wittgenstein's ([1953] 1958, §§ 184-202) 'rule-following considerations'. Kripke famously argues that

truth-conditions. What makes an utterance correct is whether it states a fact; whether what one says is the case in fact *is* the case (Hattiangadi 2007). Meaning implies correctness alright, but correctness is nonnormative.

We are assuming that we cannot understand something as a meaningful noise unless it is assessable as correctly or incorrectly uttered. Assuming this conceptual implication we want to know how to understand *correctness*. Is it an essentially *normative* notion, and if correctness is normative, in what sense of 'normative'? To answer this, I first rehearse the normativist position (sect. 2), and then consider anti-normativist objections (sect. 3). I am not interested in defending either side. The aim is rather to show that a failure to distinguish between the normativity of meaning and the normativity of use is responsible for confusion on both sides. Anti-normativists are sensitive to this distinction but, I argue, draw the wrong consequences from it. The result is that although anti-normativists are right in that correctness of use is not essentially normative, normativism can be developed into the thesis that without norms neither correctness nor incorrectness makes sense. I call this *fundamental normativism* (sect. 4). The idea is that the fact that words have meaning does not conceptually imply that one ought or may (not) use them in any particular manner; but, the fact that words have meaning conceptually implies commitments in use, although this does not mean that one ought or may so commit or that so committing conceptually implies any obligations of permissions to use words in any manner.

Fundamental normativism is Brandomian in spirit. I discuss the similarities between Brandom's view and mine (sect. 5) and defend his view against a recent anti-normativist attack (sect. 6).

On the submitted account one may consistently combine anti-

---

the question what an expression means is to be answered by a statement as to how one should apply it rather than as to how one *will* apply it (1982, p. 37). Many philosophers have fallen in line with this normativist idea (Boghossian 1989; Brandom 1998, 2001; Gibbard 1994, 2003; Whiting 2008; Wright 1980). Whether this is a correct reading of Wittgenstein's view on rule-following can be discussed (see Glüer and Wikforss 2010).

normativism, insofar as this amounts to rejecting that oughts can be directly derived from meaning, with normativism, insofar as this amounts to accepting that commitment in use is a precondition for meaning.

2. Normativism

In its simplest, normativism takes the following positive prescriptive form. Assuming that a speaker *S* means *F* by *w*,

(1) $\forall(x)$: if *x* is *F*, *S* ought to apply *w* to *x*.[3]

The obvious problem with (1) is that the fact that something instantiates properties qualifying it as a dog, quasar, horse, or whatever, does not seem sufficient for an obligation to say anything about it (Hattiangadi 2006, pp. 224-26). For instance, assume that *S* means *horse* by 'horse' and that there are horses on Gliese 581c. Failure or omission of *S* to apply 'horse' to horses on Gliese 581c would by (1) be a violation of an obligation. Generalized, (1) logically implies that for every *x*, if one's vocabulary contains a term *w* for *x*, one ought to apply *w* to *x*. Failure or omission in this is doing something wrong, as conceptually implied by the *meaning* of *w*. But the mere fact that *w* is in one's vocabulary, that one *can* apply *w* in the presence of *x*s, does not imply that one *ought*.

The dilemma can instructively be put in analogy to belief normativism. It is sometimes claimed that one ought to believe *p* if and only if *p* (Engel 2001, 2007; Shah and Velleman 2005). This logically implies that if *p*, one ought to believe *p* (Gibbard 2005). But then one ought to believe *everything* that is true. Suppose that the shape of the universe is that of a dodecahedral sphere. According to (1), that the universe has this shape implies that you ought to believe that that is its shape. If there are six thousand eight hundred and fifty seven feathers in your bedspread, you

[3] The formulations of normativist principles I use follow roughly Hattiangadi (2006) and Whiting (2007).

ought to believe that there are six thousand eight hundred and fifty seven feathers in your bedspread. But, intuitively, whatever the shape of the universe and whatever the number of feathers seems insufficient for one to be *obliged* to believe anything about that shape or that number. This is one reason for belief normativists to sophisticate their claim (Engel 2013, p. 618).[4] The case is similar for (1). Meaning normativism should be sophisticated.

Normativists agree (Boghossian 2003, pp. 36-7; Whiting 2013, MS p. 6). Sophistication is suggested either by reversing the order of implication of (1), thus

(2) $\forall$ (x): *S* ought to apply *w* to *x* only if *x* is *F*.

or by substituting prescriptive for proscriptive form, thus

(3) $\forall$ (x): if *x* is *not F*, then S ought *not* apply *w* to *x*.

In contrast to (1), (2) does not have the consequence that for all *x*, if *x* is *F* and *S* means *F* by *w*, then *S* ought to apply *w* to *x*. The idea is that the reverse implication holds. Understanding that *w* means *F* implies understanding that *S* ought to apply *w* to *x* only if *x* is *F*. So on this version we do not face the problem that if one's vocabulary contains a term *w* for *F*s then one ought to apply *w* to *everything* that is *F*.

---

[4] Boghossian argues (2003, p. 37) that this must mean that if correctly believing is normative, it cannot be normative in the prescriptive sense. He suggests that all correctness facts are normative facts but not all normative facts entail prescriptions. Thus for every *p*, if *p*, it is correct to believe that *p*, but it does not follow that you ought to believe that *p*. However, believing *p* in this case would be normatively right.

I will have more to say about this later. For now, it should be noted that Boghossian's point is easily misinterpreted. Glüer and Wikforss (2009, p. 35) say that Boghossian argues that 'correctness is a normative matter, involving "oughts"' (cf. 2014, pp. 122-23; Whiting 2009, p. 537). This is false if ascribing to Boghossian the claim that correctness implies oughts. Boghossian says no such thing.

Likewise, in contrast to (1), (3) avoids the implausible consequence that one ought to apply *w* to *x* for every *x* that is *F*. On (3) it is rather the case that *w* meaning *F* conceptually implies that applying *w* to an *x* that is *not F* is wrong, not that omission to apply *w* to an *x* that is *F* is wrong. So if *S* understands that *w* means *F*, then *S* understands that she ought not apply *w* to anything *not F*.

Sophistication can also be achieved by substituting the prescriptivity of (1) for permissibility (Whiting 2013, MS p. 6), thus

(4) $\forall(x)$: if *x* is *F*, *S* may apply *w* to *x*.

In contrast to (1), (4) does not say that a norm is violated out of omission of applying *w*, meaning *F*, to some *x* that is *F*. Not doing what one is permitted to is not equivalent to doing something one ought not, why not applying *w* to *x* if *x* is *F* is not a violation of (4). And, in contrast to (2) and (3), (4) does not imply that if *S* understands that *w* means *F*, then *S* understands that she ought or ought not apply *w* only if *x* is or is not *F*; only that she is permitted to do it if *x* is *F*. A connection between (2), (3) and (4) is brought out in the following formulation:

(5) $\forall(x)$: *unless x* is *F*, *S* ought not apply *w* to *x*,

which means that *S* is *permitted* to apply *x* to *F* only if *x* is *F* and *ought not* to do so if *x* is *not F*.

Problematic consequences of (1), then, can be responded to by certain substitutional manoeuvres either in the direction of implication, as in (2), or among prescriptives, proscriptives and permissives, as in (3) and (4), or both, as in (5). Normativism is sometimes expressed as the claim that meaning is "fraught with ought" (Gibbard 2003). This claim, properly sophisticated, appears less problematic than at first blush.

Before moving to an alternative normativist thesis, lets consider typical anti-normativist objections. This provides a context in which the alternative to be developed can be truly appreciated as an alternative.

3. Anti-normativism

Recall that normativists and anti-normativists alike accept the meaning platitude. Both sides accept that meaning conceptually implies correctness. The point of disagreement is over how to understand correctness: as essentially normative or not. Anti-normativists think they can make sense of correctness without appeal to norms.

To begin with, anti-normativism is not the claim that it is never the case that correctness implies norms. The anti-normativist distinguishes between *categorical* and *hypothetical* norms (Hattiangadi 2006, p. 228). Anti-normativists agree that meaning is hypothetically normative in that if you desire, e.g., to speak the truth, then you ought to say that it is raining only if it is. But this is not to admit that meaning is categorically normative. It is not to admit that if some expression *E* means *it is raining*, it follows that you ought (not) or may utter *E* only if it is raining. Something must be added: e.g., a desire to speak truly (Wikforss 2001). It is conditional on such contingent pragmatic considerations that correctness can be understood as normative; semantics alone is nonnormative. This means that although use may be normative contingent on pragmatic context (e.g., socially imposed conversational rules, means-ends considerations), semantic correctness does not conceptually imply norms.

Glüer and Wikforss make the point by way of arguing that, basically, using language is a practice of semantic categorization. Semantic categorizing is like sorting objects into *greens* and *non-greens* by arbitrary labelling things 'green' and 'non-green' (Glüer and Wikforss 2009, p. 36; Cf. Glüer 2001, p. 60). Insofar one manages to apply 'green' only to *greens* one categorizes correctly in the *descriptive* sense of *truly*, or in the sense of saying something that corresponds to how things in fact are (Hattiangadi 2007). There is no additional *normativity* intrinsic to semantic categorization. However, norms can always be contingently *imposed* on semantic categorization. Glüer and Wikforss put it thus,

> Of course, saying that [semantic] categorization is non-normative is not

> the same as saying that it cannot be used to derive normative consequences. Indeed, *any* categorization can be used to derive normative consequences. But not directly. Any categorization of things into *A*s and *non-A*s … can be used to derive normative consequences *if a suitable norm is in force*. (2009, pp. 36-7)

Semantic correctness, anti-normativists argue, is basically a *descriptive* term, albeit, in the context of pragmatic considerations, contingent norms may be in force. So, saying that *S* correctly applies *w*, meaning *F*, to some *x*, just means that *x* is *F*; that *S* has said something *true*. Truth and correspondence is what accounts for semantic correctness, why the correctness involved in the meaning platitude does not conceptually imply norms. This is why *meaning* is nonnormative in the sense of involving oughts (Hattiangadi 2009, p. 55; Glüer and Wikforss 2009, p. 36), although *use might be*. Given that we can separate meaning and use in this manner, we can understand the correctness of, e.g., "It is raining", with its standard meaning, in nonnormative terms: in terms of whether or not it is raining. Moreover, we can understand the normativity of the utterance as a property of its contexts of use—in terms of whether it is raining or not—by reference, e.g., to maxims such as Grice's (1989), to not say that for which one lacks evidence, to be perspicuous, and so on. In this way the anti-normativist furnishes us with an understanding of the correctness of *meaning* as nonnormative while also accounting for the contingent normativity of correctness of *use*. Hence, Glüer and Wikforss conclude, 'that *semantic* correctness is normative, is no conceptual truth' (2009, p. 36).

A final line of objection to normativism is the following. How can it follow from the fact that *w* is used by *S* to mean *F* that *S* ought (not) or may apply *w* to *x* only if *x* is *F*? After all, when we say that *S* means *F* by *w* we seem to be describing *S*'s use; how *S* is in fact using *w*. But can we from this *is*-statement derive any *ought*-statement about *S*'s use of *w*? According to Hume's Law (1968, p. 469; Cf. Hattiangadi 2007, pp. 52-3) the trouble with this is that in deriving what ought to be the case from what is the case we *introduce* a relation that has no basis in the premises. In the present context, it seems that we can describe *S*'s use of *w* and from this infer,

perhaps, something about how *S* likely *will* use *w* in the future. But as soon as we introduce the notion that *S*, by virtue of how she *is* using *w*, *ought* to use *w* in any particular manner, we make a leap to a conclusion not supported by the premises. We are smuggling in an is-ought relation between facts about past use and norms for future use. That *S* is in fact using *w* to mean *F* is, considered in abstraction from any norm in force, e.g., an agreement on how to use *w*, not something from which a statement about how *S* ought to use *w* can be derived. This being so suggests that neither the meaning of *w* nor *S*'s use of *w* suffices for the derivation of any of the normativist claims (1)-(5) considered above. As per Glüer and Wikforss's point (2009, pp. 36-7), it seems that normativity must be exogenous to both meaning and facts about use; it must be imposed and contingent.

Now, I am neither particularly interested in whether anti-normativists have a knockdown argument, nor in whether normativists have a solid defence. What is interesting about this debate is that, it seems to me, one can quite consistently accept the claims made on both sides with only slight modifications to the normativist position. This is what I turn to next.

4. Fundamental normativism

Recall the anti-normativist claim that semantic categorization, like sorting objects into *greens* and *non-greens*, is nonnormative (Glüer and Wikforss 2009). The correctness involved in saying that something is green is determined by whether it *is* green, by whether one would say something true, not by it in addition being prescribed or permitted to say that it is green. Lets assume that this is right.

I argue in this section that semantic categorization cannot be understood nonnormatively. I will not argue that oughts can be directly derived from meaning though. This means that I accept both the anti-normativist concerns and accept that meaning is normative. If this is possible then it is possible to understand meaning as normative in a sense that is perpendicular to the two positions considered above.

To see that this is possible consider first that *categorizing*, as an activity, is not true or false. It is something we *do*—an ability or practice. If correctness, which we are assuming is essential for meaning, is, pace anti-normativists, basically a matter of categorization, then meaning conceptually implies a practice which is not qua practice true or false. What is true or false is that a practice of categorization is or is not engaged—that we do or do not categorize. Assuming the meaning platitude and anti-normativism, the question then is: What about a practice of categorization qualifies it as *semantic* categorization? This is important because unless the categorizing is semantic, correctness conditions do not apply, and if correctness conditions do not apply then the categorizing labels, e.g., 'green' and 'non-green', have no *meaning*. Hence, we must account for what qualifies categorization as semantic if we want to account for correctness, which we are assuming together with anti-normativists and normativist alike is key to distinguish noise with meaning from noise without.

Now, note further that the anti-normativist claim that the correctness of an expression is determined by it being true or false presupposes that the expression is semantically contentful; '*x* is green' is an expression with semantic content only if it is correct to utter it. And it is correct to utter it only if *x is green*. Given that we are now asking how it is that an expression can be part of a practice of semantic categorization, it should be obvious that an appeal to it being an expression susceptible to anti-normativist correctness conditions is *circular*. Saying that 'green' is correctly applied to *x* only if *x* is *green* is to presuppose that 'green' means *green*, which is just to presuppose that it is part of a practice of semantic categorization. What we want to know is under what conditions it can be said that an expression is part of such a practice such that the expression can be said to have meaning. What is required is to account for what it is about arbitrary labelling things, e.g., *greens* and *non-greens*, that bestows semantic content on the labels, *ab initio* as it were.[5] What is a categorizing

[5] Part and parcel of this might be put in Sellaresque context: 'if the ability to recognize that x looks green presupposes the concept of *being green*' (1956, §

practice such that the categorizing involves meaning? These considerations conspire to show that in accounting for what makes practices of categorization semantic an appeal to truth-conditions is circular. This means that we cannot appeal to the anti-normativist idea that correctness is determined by truth-conditions in attempting to account for semantic categorization.

It is instructive in attempting to formulate the sought after account to consider two insidiously similar claims that meaning is normative. This will show that such claims are hostage to a subtle confusion. On the one hand meaning is sometimes claimed to be normative in the sense that practices of *attributing* meaning is normative. On the other hand it is sometimes claimed that *meaning* attributed is normative (Gibbard 1994). In the first sense the claim is that saying that *S* means *F* by *w* is, as Gibbard puts it, to be 'speaking oughts' (2003, p. 85). Attributing meaning to *S*'s expression just is saying that *S* ought to apply it only in certain circumstances (e.g., only if *x* is *F*). In the second sense the idea is that if *w* means *F*, then the fact that *w* means *F* implies oughts. But now it is the *meaning* of *w* that is normative (Whiting 2007; 2009), not the *attribution* of meaning to *S*'s *use* of *w*. Before pointing out the importance of avoiding this confusion lets consider another domain where it is often made. I am thinking of claims that belief is normative.

On the one hand, the claim that belief is normative is sometimes formulated as it being constitutive of understanding *what it is for S to believe p* or to *attribute* a belief that *p* to *S* that one understands that *S* ought to believe *p* only in certain circumstances (e.g., only if *p*). Here it is the conceptual role of 'belief' that is normative. Having the concept involves understanding that oughts apply (Boghossian 2003). On the other hand, the claim is sometimes put as it being constitutive of *S's* belief *that p* that *S* ought to have it only if *p*. In the latter case, it is not the conceptual

---

19)—if sorting things into *greens* and *non-greens* qualifies as semantic categorization only by presupposing conceptual content—then 'it would seem that one couldn't form the concept of *being green*, and, by parity of reasoning, of other colors, unless [one] already had them' (Ibid).

role of 'belief' that is normative, but *contents* believed (Ibid).[6]

In light of these observations and our assuming anti-normativism, it emerges that we have to distinguish the semantic normativity of *meaning* from the pragmatic normativity of *use*. Considering the above discussion of normativism and anti-normativism it seems that what we have been focusing on this far is the normativity of *use*; i.e., pragmatic norms. After all, the debate has focused on whether, *assuming that S means F by w*,[7] *S* is permitted or obliged (not) to use *w* in certain contexts. But this means that the fundamental question, whether *w meaning F* conceptually implies norms, is not addressed. It is assumed already that *S means F* by *w*, and then asked whether on this assumption *S* is permitted or obliged (not) to certain uses of *w*. Clearly, then, the question about the normativity of

---

[6] Boghossian oscillates violently between three formulations:

> Does the fact that such [content] *attributions* are normative reveal something normative about our *notion* of content, or does it reveal something, rather, about our *notion* of belief? Do we have here a thesis of the normativity of *content*, or a thesis of the normativity of *belief*? (2003, pp. 39-40, emphases added)

Clearly, Boghossian should not move, first, from *practices of attributing* content (belief) being normative to *the notion* 'content' ('belief') being normative, and then to *content (belief)* being normative (Cf. 2008, pp. 212-13). There are at least two implicit transitions involved in this reasoning. Although there may be some connection—I do not see clearly what it is though—questions about norms for practices of attribution should be kept apart from questions about norms for concept possession and application, which should be kept apart from questions about content norms. Boghossian goes on asking all three questions as if they were one, expecting, it seems, that an answer to one just is the same as an answer to all three (2003, p. 41). It is then far from clear what meaning (content) normativism really amounts to. Is it *the notion* of meaning (content) that is normative (2008, p. 213) or is it *meaning* (*content*) that is normative (Ibid, p. 214), or is it that the *practice of attributing* meaning (content) is normative (Ibid, p. 213)?

[7] Consider all normativist principles (1)-(5) above. We have considered each as follows: 'Assume that *S* means *F* by *w*…' We have then argued that the meaning of *w* is nonnormative. In a sense we have been assuming all along that *S* is already engaged in a practice of 'semantic categorization'. Of course, whatever *then* is or is not normative is not meaning but what we *do* with language.

meaning has all along taken a backseat position with respect to the question whether, *given its meaning*, using *w* is normative. If we want to know whether meaning is normative this strategy will not get us to the bottom of our inquiry. This is precisely why I propose that we focus analysis precisely on the *semantic* in 'semantic categorization'. What is it about a practice of categorization that qualifies it as semantic, such that we can understand it as subject to correctness (normatively or nonnormatively understood), hence as involving meaning?

Here is a first stab. Assume that 'horse' means *horse*. Assume also with anti-normativists that the fact that 'horse' means *horse* does not conceptually imply, e.g., that Zorro ought (not) or may say that Tornado is a horse only if Tornado is (not) a horse. If Zorro says, "Tornado is a horse" then Zorro is speaking truly and if he says, "Tornado is a soldier of the Spanish colonial army" then he is speaking falsely. Insofar Zorro's use of 'horse' is concerned, i.e., *given* that 'horse' means *horse*, Zorro may do whatever he wants with 'horse'.

Suppose now that Esmeralda tells Zorro that his horse is waiting for him in the alley. However, Esmeralda means by 'horse' *soldiers of the Spanish colonial army*. Zorro jumps from the balcony into the alley and lands, not in the saddle, but encircled by Spanish colonial soldiers. Esmeralda was speaking incorrectly. According to our assumptions, semantic correctness is truth-conditional. Esmeralda was speaking truly, hence correctly, by *her* assessment but falsely, hence incorrectly, by Zorro's. So far we might agree with the anti-normativist.

Suppose next that Esmeralda's use of 'horse' diverges extensively from Zorro's. Zorro may still learn, over time and through radical and charitable interpretation (Davidson 1973, p. 18; 1984, pp. 15-16) what Esmeralda means by 'horse', assuming that she means something by it. Zorro might attribute to Esmeralda a pattern of linguistic behaviour such that she can at least be held to conform to his 'inveterate' habit of speaking about objects (Quine 1957), and do so mostly correctly (Davidson 1990, pp. 319-20). Here it is important to consider what Zorro is *doing*.

Zorro is treating Esmeralda as speaking about objects coherently and as getting things right most of the time, and he takes himself to be entitled

to attribute such properties to Esmeralda's linguistic behaviour. He takes Esmeralda to be engaged in 'practices of categorization' that entitle her to be charitably interpreted. Now, if he does not do this then all bets are off and he might just as well refrain from assuming that Esmeralda is engaged in *semantic* categorization; for then he would not recognize her as applying concepts or making meaningful noise at all, hence recognize no correctness or incorrectness in what she does. Of course this refusal might be elicited for several reasons. The most obvious reason is if Esmeralda strikes Zorro as not committed to any modicum of coherence in her categorization. This, then, seems to be fundamental for some practice to *be* assessable as semantic categorization: A practice of categorization can be assessed as semantic, hence as correct or incorrect, if and only if the categorizer can be counted on as committed to a minimal standard of coherence, thereby entitling its hearers to interpret it (charitably) as so committed. No interpretation is possible in abstraction from the interpreter recognizing the interpreted as exhibiting at least minimal coherence in linguistic behaviour.

Considering the Glüer and Wikforss-argument it seems that to understand *S* as *semantically* categorizing greens and non-greens—to be engaged in a practice to which correctness apply—is to understand *S* as using 'green' and 'non-green' in a minimally coherent manner. If this conceptual implication does not hold we cannot assume that *S* means anything with 'green', for we do not then have any criterion against which *S* can be said to categorize greens correctly or not. But, now, is *coherence* essentially normative, such that in understanding someone as meaning something with her expressions we are, necessarily, understanding her as subject to obligations and permissions?

Coherence, it seems, is *not* essentially normative. For instance, my fridge has given off a murmuring noise every morning at 9:15 a.m. for about five minutes for the last six years. I understand this as a coherent noisemaking, but I do not understand the fridge as obliged or permitted to murmur every morning between 9:15 a.m. and 9:20 a.m. Nor do I think that the fridge means anything by murmuring, although I understand it as the cooling system being turned on. Coherence, it seems, is not normative. Moreover, coherent noisemaking is not sufficient for *meaning*, lest we are

committed to say that, e.g., my fridge *means* and that its cooling system is turned on. Coherence, then, is neither normative nor sufficient for semantic categorizing.

Perhaps what is missing is *reliable differential noisemaking* in response to certain stimuli, like greens and non-greens. But this will not distinguish semantic from non-semantic categorization either. Consider the concept 'cold'. A thermostat can be relied on to make noise whenever it gets too cold. Although thermostats are reliable differential noisemakers in response to temperature, thermostats do not *mean* it is cold by emitting <beep>. Thermostat alarms do not seem to qualify as concept applications. Here we have an instance of coherent and reliable differential noisemaking, like that of categorizing objects into greens and non-greens, but the categorizing is not semantic. Thermostat alarms have no meaning such that were a thermostat to get things wrong too often we would be committed to interpret it as meaning something different by its noises in order to make it intelligible. To quote Brandom:

> Merely reliably responding differentially to red things is not yet to be *aware* of them *as* red. Discrimination by producing repeatable responses (as a machine or a pigeon might do) sorts the eliciting stimuli, and in that sense classifies them. … (If instead of teaching a pigeon to peck one button rather than another under appropriate sensory stimulation, we teach the parrot to utter one noise rather than another, we get only the vocal, not yet the verbal.) (2001, p. 17)

Analogously, merely 'sorting objects into greens and non-greens', even if a coherent, reliable categorizing noisemaking, is not yet noise with meaning, assessable as correctly or incorrectly emitted; it is vibrations and stretching, etc., of vocal cords. Saying that *S* reliably categorizes stimuli by being disposed to coherently emit one noise and not another in response to it is not sufficient for saying that *S* is engaged in a practice of semantic categorization. Anti-normativists successfully argue, we are assuming, that the fact that 'cold' means *cold* does not conceptually imply obligations or permissions. In assuming this, though, we are only assuming that, *given* the meaning of 'cold', no norms can be directly derived regarding *use*. But, to

repeat, what about categorization accounts for its semantics, such that we can even begin to make sense of 'green' meaning *green* and 'cold' meaning *cold*, even if we accept that *using* 'green' and 'cold' is not, assuming their meaning, essentially normative? Being disposed to coherent and reliable differential noisemaking is not sufficient.

Here is a second attempt. Semantic categorization conceptually presupposes a (perhaps implicit) practice of *committing* to coherent and reliable differential noisemaking. Being committed, not merely responding to stimuli with a brute sounding-off, is what separates speakers from mere coherent, reliable differential noisemakers, such that only the former can be assessed as subject to standards of correctness. By committing to coherent, reliable noisemaking one undertakes a role (in a linguistic community, say) as someone responsive to demands for correction (Brandom 2001, p. 190). One can be relied upon, not only as causally responding coherently to eliciting stimuli as a machine or pigeon might do but as also entitled, and entitling others, to respond similarly in similar circumstances by deference (Brandom 1998, p. 122). In this respect the coherent, reliable categorizing of faces and non-faces of the autofocus function of a camera is quite different; a difference between categorization that does *not* and categorization that *does* imply semantic correctness. Of course, once the latter is established—once semantics is attached to categorizing practices—one can do whatever one wants with 'green', 'face' or 'horse' (*pace* Glüer and Wikforss 2009)—e.g., use it in metaphor—without thereby violating any categorical obligation conceptually implied by their meaning. We can also envisage departures from pragmatic, contingent norms insofar interpretation can restore intelligibility. One may depart radically from pragmatic hypothetical maxims—e.g., from Gricean maxims—precisely to get metaphor and irony across to an audience. What is unintelligible, on the present proposal, is to understand categorization as semantic if the noisemaker, like a thermostat, is not responsive to commitments and as not entitling others to demand correction.

Obligations and permissions can on this approach to be understood as (attempts at) *expressing* what we are *doing* when we semantically categorize. Thus, saying that *S* ought to or may (not) say that *x* is *F* only if

x is *F* is to make explicit commitments implicit in *S*'s practice of categorizing *F*s and *non-F*s. Normative vocabulary emerges as a pragmatically mediated meta-vocabulary for *speaking about what one must do in order to count as a speaker* (Brandon 2008, pp. 12-13, 110-11). 'Speaking oughts'—indeed, speaking at all—conceptually implies a practice of committing to coherent and reliably differential noisemaking. This practice of committing is what allows others to hold one to standards of *correctness*, to rely on one as responsive to demands for giving reasons for claims and as responsive to demands for correction. For instance, if *S* says '*x* is green', which *S* may be entitled to by observing *x*, then *S* is committed to '*x* is coloured' (if considering or being asked about the latter). If *S* cannot be understood as responsive to entitlements-commitments in categorizing, e.g., greens and non-greens, then *S*'s listeners cannot understand *S* as meaning *green* by 'green'. And, if this is the case for *S*'s categorizing in general then *S* cannot be understood as speaking. What this amounts to is an understanding of meaning as conceptually presupposing a normative structure of commitment-entitlement relations internal to practices of noisemaking such that noisemaking can count as semantic. It also amounts to admitting that *given* its content a concept may be applied in whatever way one wants (there is no categorical norm as those in (1)-(5)) insofar one is responsive to demands for correction.

It can be seen, then, that one may consistently concede the anti-normativist arguments and insist that, still, meaning is normative. This kind of normativism is not of the generic stripe that one ought or may (not) use *w* in particular ways *given* its meaning. Nor does it imply the incorrectness of such a view. As stated, I am quite uninterested in defending either side of debate as it now stands. Being agnostic as to the present normativist–anti-normativist debate we can still and should understand meaning as normative. If we do not, then we cannot make sense of a practice of categorization as a practice of semantic categorization. I call this view *fundamental normativism*: sematic content relies on a normative pragmatics implicit in categorization.

I assume the reader will see a familiarity between fundamental

normativism and Brandom's (1998) view. Let me therefore elaborate on particular similarities and differences between the two.

5. Brandom and fundamental normativism

One thing to be noted at the outset, regardless of the siblinghood between Brandom's view and mine, is that the argument of the previous section has been largely neglected in the normativist–anti-normativist debate. An exception is Glüer and Wikforss's (2009, pp. 60-62) criticism of Brandom, which I will discuss shortly. At this point I merely want it be known that the present argument is quite novel to the debate considered above, regardless of its own merits.

I take it that Brandom holds the view that normative properties implicit in practices that give content to concepts are such that oughts apply. For instance, he argues that

> we can understand making a claim as taking up a particular normative stance towards an inferentially articulated content. It is *endorsing* it, taking *responsibility* for it, *committing* oneself to it … whereby undertaking one commitment *rationally* obliges one to undertake others, related to it as inferential consequences. (2008, p. 113)

According to Brandom, the claim "This is red" *obliges* to certain other claims that it commitment-entails, e.g., "This is coloured", and it is entitlement-entailed by certain other claims, e.g., "This is scarlet", such that one *may* (in the deontic sense) say "This is red" if one is entitled to "This is scarlet" and one *ought* to say "This is coloured" if one is committed to "This is red".

Here I disagree with Brandom. Although I argue that commitment-entitlement relations internal to practices of categorization are necessary and sufficient for terms employed for categorizing to have meaning, I do not believe that it follows from the necessary and sufficient conditions being satisfied that there is anything a speaker *ought* or *may* do. Such obligations and permissions I have argued are better understood as

hypothetical. One of my reasons for this was that it seems to better capture the sense in which it can be entirely *appropriate* to *violate* the relevant norms in the context of a conversation in which one wants, e.g., to speak in metaphor or tell a joke. These are situations in which by saying something that one *ought not* say one can convey a message to one's listeners different from what one would convey if adhering strictly to one's position in entitlement-commitment relations. In doing so one is not doing something wrong but something warranted by the pragmatics of the conversation.

Another reason why I do not follow Brandom in this regard is that I conceive of the implicit normative structure of discursive practices somewhat differently than he. Brandom's idea, which he draws from Kant and Sellars, is that

> normative vocabulary (of which 'ought' and 'should' are paradigmatic) has the logical expressive function of making explicit in the form of something that can be *said* (put in the form of a claim) an attitude that otherwise could be implicit only in what is *done*—namely, the endorsement of a pattern of practical reasoning. (1998, p. 271)

To me, this idea of an explicitating logical function of normative vocabulary as allowing us to *say* what we *do* such that we can qualify as speakers to begin with (cf. 2001, p. 89) is too narrow. To see why, consider Glüer and Wikforss's point that *any* categorization of things into *A*s and *non-A*s can be used to derive normative consequences if a suitable norm is in force; e.g., by being socially imposed. I take it that normative vocabulary is not just a meta-vocabulary for explicitating normative (reasons-providing) relations implicit in practice. It also (perhaps more importantly) has the pragmatic declarative function of *imposing* in the form of something that can be said normative (reasons-providing) relations *not* already implicit in practice, and hence *not* something the vocabulary is making or can make explicit.

Searle (1995) presents several examples of this sort for social and institutional kind concepts. For instance, a speaker might claim that the claim "*X* is president" commitment-entails "*X* is commander in chief". This

speaker might not be making explicit normative relations implicit in applications of the concepts 'president' and 'commander in chief' but to *make it the case* that his audience accepts the commitment-entailment suggested, in which case "*X* is president" and "*X* is commander in chief" *will be* commitment-preserving. Since, in this case, the concepts were not inferentially related in accordance with norms implicit in their use prior to the declaration, the speaker did not make any such norm explicit. What he did, if successful, was to impose a *new* commitment-entailment from 'president' to 'commander in chief' such that their contents are now inferentially related. If, now, someone is demanded a reason for saying "*X* is commander in chief" he may point out that he is entitled to "*X* is president", from which "*X* is commander in chief" can be entitlement-derived. What all this means is that normative vocabulary is not in the business merely of explicitating implicit normative relations in use, but also in the business of imposing, in the form of declarations, what normative relations should be in force.

Expanding the pragmatic function of normative vocabulary in this manner is not, as far as I can see, in any way incompatible with Brandom's view. Nor, of course, does it imply the falsity of his statement in the quoted passage. It does, however, reach out to the anti-normativist suspicion that what we are really saying when we say that one *ought* or *may (not)* so-and-so is that normative relations have been *declared* to be in force. I also believe that recognizing this richer pragmatic function of normative vocabulary supports Brandom's general project by implying that the material inferential relations, in which expressions are embedded and that bestow them with content, are essentially *contestable*.

To see what I mean, consider Brandom's example (2001, 69-70) of the concept 'Boche' (borrowed from Dummett 1973). This example is meant to capture a process of conceptual change and its implication for commitment-entitlement relations. 'Boche' is introduced as entitlement-*entailed* by 'German'. Thus, if someone is of German nationality one is permitted to say that he is Boche. However, 'Boche' also entitlement-*entails* 'barbarous'. So, if one is permitted to say that someone is Boche then one is entitled to say that he is barbarous. By transitivity, the material

inference from '*X* is German' to '*X* is barbarous' is entitlement-preserving. Denying that there are Boches is to deny that there are Germans; admitting that there are Boches is to accept that Germans are barbarous. Thus, if one refuses the implication from German nationality to barbarousness then one has to refuse to use 'Boche' because of the implicit material commitment-entitlement inferential relations its introduction into one's vocabulary imply.[8] Concerning the explicitating function of normative vocabulary, Brandom writes:

> The proper question to ask in evaluating the introduction and evolution of a concept is not whether the inference embodied is one that is already endorsed, so that no new content is really involved, but rather whether that inference is one that *ought* to be endorsed. The problem with 'Boche' and 'nigger' is not that once we explicitly confront the material inferential commitment that gives the term its content it turns out to be novel, but that it can then be seen to be indefensible—a commitment we cannot become entitled to. We want to be aware of the inferential commitments our concepts involve, to be able to make them explicit, and to be able to justify them. (2001, pp. 71-2)

It is precisely in the context of discussions like these, I argue, that the logical expressive function of normative vocabulary, emphasized by anti-normativists, of *declaring* the appropriateness of certain material inferential relations as *being in force* should be recognized. The idea is that many concepts might *not* commitment-or-entitlement-entail material inferences to certain contents but be declared to do so by means of declaratives like "*X* is to count as *Y*" ("Germans are to count as barbarous") from which it would follow that a consequence of application of '*X*' obliges an inference to '*Y*' even if that inference is not already implicit in '*X*'. This means that there may be several concepts whose contents are not determined by material inferential relations in which they are implicitly embedded but whose contents are rather socially imposed. And the importance of this

[8] Similar concepts are 'nigger', 'whore', 'faggot', 'communist', etc. The point is that these concepts are caught up in material inferential relations whose circumstances of application are descriptive (e.g., 'German') but whose consequences of application are evaluative (e.g., barbarous).

point is that it might actually *not* be the case that one can derive from the normative inferential relations in which concepts stand that one *ought* apply them. If the 'oughts' are imposed then they are hypothetical, such that it is only conditional on one's accepting the material inferences they commit one to that one ought to apply them.

Thus, I agree with Brandom that semantic content conceptually presupposes responsiveness to commitment-entitlement relations. But I disagree with the view that these relations involve categorical obligations for concept application because, first, in some pragmatic contexts, e.g., metaphor, it might actually be precisely by not saying what one ought that the meaning of what one says is determined and, secondly, in many cases commitment-entitlement relations are socially imposed and therefore not such that one ought, in the *categorical* sense, conform to the norm. Oughts are rather, I'd like to say, *essentially contestable*. By this I mean that although all conceptual content conceptually presupposes commitment-entitlement relation, there is *no* concept embedded in commitment-entitlement relations that *obliges* one to certain inferences. Fundamental normativism, in this context, can be expressed as the view that all semantic content conceptually presupposes implicit normative relations of commitment-entitlement in use, and no commitments-entitlement relations are such that one ought categorically to conform with them because all such relations are essentially contestable (recall Brandom's own example with 'Boche').

As I see it, content and meaning is fundamentally normative in that concept possession and application is inevitably a matter of bowing to or turning against, accepting or rejecting—essentially, taking a stance on—content-constitutive normative relations. Possessing a concept *is* (implicitly) taking a normative stance in a field of normative material inferential relations; and concept use *is* a (implicitly) normative action in that field, of committing to and entitling others to material inferences between contents of concepts used. Sometimes normative relations not implicit in possession or use are not evolved through (fieldwork) normative action but imposed from the sidelines, as it were, in attempts to change the game of what counts as correct. What this picture, of concept possession

and use as essentially normative stances and actions in a field of material commitment-entitlement inferential relations, where some inferential relations are imposed, shows is that all *oughts* in play are *contingent*. Playing the game is a necessary and sufficient condition for counting as having and using concepts but one cannot directly derive categorical oughts from playing the game because oughts are contestable.

I promised earlier to discuss Glüer and Wikforss's criticism of Brandom's understanding of content as normative. Defending his view, with the material added by fundamental normativism, against their criticism is my aim in what remains.

6. The threat of regress

According to Glüer and Wikforss, Brandom's normativism runs a vicious regress. The dilemma is supposed to be that correctness conditions are, on Brandom's view (1994, p. 25), instituted in attitudes of 'taking something to be correct-according-to-a-practice'. Glüer and Wikforss (2009, pp. 61-2) note that these attitudes are themselves essentially normative; i.e., something can correctly be taken to be correct according to a practice. But, surely, what is correctly taken to be correct according to a practice might or might not be correctly so taken. And so on for all attitudes of taking something to be correct according to a practice. Glüer and Wikforss conclude, drawing from Gideon Rosen's (1997, pp. 167-68) criticism of Brandom, that the view

> does not provide any insight into how any one of these implicit norms actually is 'instituted', or made, by us. To provide such insight, we would, at some stage of the hierarchy, need to be told *in non-normative terms* what the norm-instituting behaviour exactly consists in. (2009, p. 62)

Contrary to Glüer and Wikforss's conclusion I argue that Brandom's approach does not run the regress in question. That it does not can be appreciated by noting that Brandom's understanding of norms is of them as *socially articulated* in the game of giving and asking fore reasons. He

writes,

> propositional contents should be understood in terms of their *social* articulation—how a propositionally contentful belief or claim can have different significance from the perspective of the individual believer or claimer, on the one hand, than it does from the perspective of one who attributes that belief or claim to the individual, on the other. The context within which concern with what is thought and talked *about* arises is the assessment of how the judgments of one individual can serve as reasons for another. The representational content of claims and the beliefs they express reflect the social dimension of the game of giving and asking for reasons. (2001, pp. 158-59)

Brandom is here arguing *not* that it is either necessary or sufficient for anyone merely to take something to be correct according to a practice, but that there is an essentially *social* dimension involved, such that content is determined by being caught up in the interpersonal practice of giving and asking for reasons. Making an assertion, e.g., is understood as playing the pragmatic role of undertaking a commitment whose content is articulated by what other assertibles it is inferable from and what other assertibles are inferable from it. This means that no one individual's attitude of taking some claim to be correct in the sense of being entitlement-or-commitment-inferable from other claims is necessary or sufficient for it be an articulation of the content of the claim. What is necessary and sufficient is that the claim serves the pragmatic function of committing the speaker to inferences for which she can be held responsible by listeners that are in position to ask the speaker for reasons for the claim made. For instance, the content of the claim "*X* is red" is socially inferentially articulated by what it commitment–entitlement-entails the speaker to in the ears of listeners for whom the claim can serve as a reason for further claims (cf. ibid, p. 192).

Now, the 'possibility of an infinite "hierarchy of critical stances"' that Glüer and Wikforss (2009, p. 62) complain about is not such that a regress threatens. Instead, commitment-entitlement inferential relations through which contents is articulated *level out* in the social dimension in which what reasons a claim gives, and what reasons can be demanded for it, is a function of the attitudes of members of the linguistic *community* in which it

is made (Brandom 1998, p. 626). In this sense, there is no and indeed need not be an 'infinite hierarchy of critical stances' or normative attitudes in order that content or meaning be determined. Instead, what is needed is a field of socially contested implicit commitment-entitlement material inferential relations in which claims are embedded. Correctness is determined in the context of the social contest over implicitly normative inferential relations between claims. We have in this context no reason to ask for, as Glüer and Wikforss do, a *further* or ultimately *final* level at which correctness is instituted. Insofar people have attitudes and assess each other's performances in terms of what follows from what—what Brandom calls 'scorekeeping' (Ibid, pp. 165-67)—and insofar this is an *on-going* contest over the field of such inferential relations, there need be no further level.

If the notion of 'an infinite hierarchy of levels' of attitudes of taking-to-be-correct is thought of in 'vertical' terms, where 'at some stage in the hierarchy' (Glüer and Wikforss 2009, p. 62) we must come to an end and say in non-normative terms what correctness consists in, then the answer is: the hierarchy levels out in the 'horizontal' field of inferential relations whose normative significance consists in people accepting or rejecting claims as providing reasons for other claims. Below this field it makes no sense to dig for conceptual content. It seems that this is precisely what Glüer and Wikforss do; they seem to suppose that access to a point of view 'beyond' or 'outside' of a community of speakers is necessary in order that we can interpret the community as making meaningful noise. It is quite unclear why this should be the case though.

Perhaps this halting of their argument will not convince anti-normativists. I am not sure what they would then be asking for. It might be that they want a reduction of the normative to a behavioural, functional, or other suitable level of description; that 'the regress above, be it ever so benign, at least indicates a serious flaw' (ibid) in Brandom's non-reductive normativism. In this case I have no quarrel with the anti-normativist's concern. But I do not see why the *possibility* of speaking about *non-normative* behaviour below the field of the social game of articulating implicitly normative commitment-entitlement inferential relations that

bestow content and meaning should imply that content and meaning is *non-normative*. If the anti-normativist point is that a story about the natural evolution of language among animals *can* be told without allusion to norms, including a story about how animals pass from mere reliable differential noise-makers to stimuli to mastering inferential relations in which claims stand that make them contentful (cf. Brandom 2001, p. 162), then their project would take us far, far afield indeed. And if such a project were successful it still would not be incompatible with the notion that to speak about meaning and content is to speak about norms (and, indeed, that to speak at all conceptually presupposes norms). *Meaning* and *content* can be fully analysed in normative terms, but of course there is much going on in thinking and speaking, e.g., neural firings and stretching of vocal cords, as investigated in the special sciences, that we can *also* speak about. It is far from clear, though, that our being able to speak about the latter should induce a conviction that, therefore, when we speak about meaning and semantic content we are certainly not speaking about norms.

7. Conclusion

We have been assuming with normativists and anti-normativists that a conceptual implication of understanding that *w* means *F* is to understand that *S*'s use of *w* can be *correct* or *incorrect*; the so-called meaning platitude. This is what distinguishes the *vocal* from the *verbal*, the coherent and reliable differential disposition to brute sounding-off in response to stimuli, on the one hand, from semantic categorization, on the other.

It emerged when we assumed that anti-normativism is correct that although the correctness involved in meaning does not conceptually imply categorical oughts in the sense that there are certain things one ought or may say in abstraction from the *imposition* of such norms, still, the very assumption that a sign or sound *can be assessed as correctly or incorrectly employed* presupposes practices of implicitly committing and entitling to material inferential relations. Taking normative stances of committing and entitling is a conceptual presupposition for a practice engaged to be

recognizably a practice of semantic categorization. Therefore, although no categorical oughts for future use can be directly derived from correctness, to understand correctness just is to understand the practice as underwritten by normative attitudes. Hence to understand a piece of vocabulary as having meaning—understanding it as part of a language—is to understand it as involving a practice of committing and entitling. It is not the case that one is categorically obliged or permitted to employ the vocabulary one way or another but that if one does not recognize the commitments and entitlements of employing it then one does not speak it.

Obligations and permissions can be understood as *codifying*, in the sense of making explicit as something that can be said, or *imposing* in the form of something declared, contingent oughts on use once a vocabulary is employed. These oughts are essentially socially contestable in the sense that once one counts as a player in the field of a language game—i.e., once one can be understood as standing responsible and in a position to hold others responsible to correction—what one ought or ought not do is not determined by some static rule which, once in play, stretches the lines along which one ought to move through the whole field. On the fundamental normativism I have elaborated, concept possession (having a position in the field) *is* a normative stance and concept use (making a move in the field) *is* a normative action. No stance or action is incontestable; each updates the score as well as what inferential moves should count as commitment-entitlement compatible or incompatible with what moves and positions. Attempts to *impose* a structure on the field by declaration are analogous to someone vying for the role of coach or referee, as being in a privileged position of authoritatively staking a path for the others. But this position is itself part of the game and is as such essentially contestable. Moreover, there is no position from beyond the sidelines from which we can speak about this game. Speaking about meaning and content is itself normative meta-manoeuvring within the game; it is like two players starting to argue over what they are doing such that they can stand and argue over what they are doing such that they can stand and argue over what they are doing… Such an argument can only ever get off the ground if the two hold each other to standards of correctness; e.g., one takes the other

to be responsible for a move that commits her to a position that she fails to take, or one takes oneself to be entitled to a position from a given move that another disputes that one is entitled to. All of this is *essentially* normative positioning and action, which institutes *contingent* directions for further play. One may consistently take the position, as I have done here, of denying that any particular directions for concept use are essential to the game and insist that to have and use concepts conceptually presupposes that one commits, and is entitled by others, to a position in the field.

References:

Boghossian, P. A. (1989). The Rule-Following Considerations. *Mind*, 98, 507–549.

Boghossian, P. A. (2003). The Normativity of Content. *Philosophical Issues*, 13, 31–45.

Boghossian, P. A. (2008). Is Meaning Normative? In *Content and Justification*. Oxford: Oxford University Press.

Boghossian, P. A. (2012). Blind Rule-Following. In A. Coliva (Ed.), *Mind, Meaning, and Knowledge: Themes from the Philosophy of Crispin Wright*. Oxford: Oxford University Press.

Brandom, R. (1998). *Making it Explicit*. Cambridge: Harvard University Press.

Brandom, R. (2001). *Articulating Reasons*. Cambridge: Harvard University Press.

Brandom, R. (2008). *Between Saying and Doing: Towards an Analytic Pragmatism*. Oxford: Oxford University Press.

Davidson, D. (1973). On the Very Idea of a Conceptual Scheme. *Proceedings and Addresses of the American Philosophical Association*, 47, 5–20.

Davidson, D. (1984). Communication and Convention. *Synthese*, 59, 3–17.

Davidson, D. (1990). The Structure and Content of Truth. *The Journal of Philosophy*, 87, 279–328.

Dummett, M. (1973). *Frege: Philosophy of Language.* New York: Harper and Row.

Engel, P. (2001). Is Truth a Norm? In P. Kotatko, P. Pagin & G. Segal (Eds.), *Interpreting Davidson*. Stanford: CSLI Publications.

Engel, P. (2007). Belief and normativity. *Disputatio*, 2, 179–203.

Engel, P. (2013). Sosa on the normativity of belief. *Philosophical Studies*, 166, 617–624.

Gibbard, A. (1994). Meaning and Normativity. *Philosophical Issues*, 5, 95–115.

Gibbard, A. (2003). Thoughts and Norms. *Philosophical Issues*, 13, 83–98.

Gibbard, A. (2005). Truth and Correct Belief. *Philosophical Issues*, 15, 338–350.

Glüer, K. (2001). Dreams and Nightmares. Conventions, Norms and Meaning in Davidson's Philosophy of Language. In P. Kotatko, P. Pagin & G. Segal (Eds.), *Interpreting Davidson*. Stanford: CSLI Publications.

Glüer, K., & Wikforss, Å. (2009). Against Content Normativity. *Mind*, 118, 31–70.

Glüer, K., & Wikforss, Å. (2010). Es braucht die Regel nicht: Wittgenstein on Rules and Meaning. In D. Whiting (Ed.), *The Later Wittgenstein on Language*. Basingstoke: Pelgrave Macmillan.

Grice, H. P. (1989). *Studies in the Ways of Words*. Cambridge: Harvard University Press.

Hattiangadi, A. (2006). Is Meaning Normative? *Mind & Language*, 21, 220–240.

Hattiangadi, A. (2007). *Oughts and Thoughts*. New York: Oxford University Press.

Hattiangadi, A. (2009). Some More Thoughts About Semantic Oughts: A Reply to Daniel Whiting. *Analysis*, 69, 54–63.

Hume, D. ([1888] 1968). *A Treatise of Human Nature*. L. A. Selby-Bigge (Ed.), Oxford: Clarendon Press.

Kripke, S. (1982). *Wittgenstein, Rules and Private Language*. Cambridge: Harvard University Press.

Quine, W. V. O. (1957). Speaking of Objects. *Proceedings and Addresses of the American Philosophical Association*, 31, 5–22.

Rosen, G. (1997). Who Makes the Rules Around Here? *Philosophy and Phenomenological Research*, LVII, 163–171.

Sellars, W. (1956). Empiricism and the Philosophy of Mind. In H. Feigl & M. Scriven (Eds.), *Minnesota Studies in the Philosophy of Science, Vol. 1*. Minneapolis: University of Minnesota Press.

Shah, N., & Velleman, J. D. (2005). Doxastic Deliberation. *The Philosophical Review*, 114, 497–534.

Whiting, D. (2007). The normativity of meaning defended. *Analysis*, 67, 133–140.

Whiting, D. (2008). On Epistemic Conceptions of Meaning: Use, Meaning and Normativity. *European Journal of Philosophy*, 17, 416–434.

Whiting, D. (2009). Is Meaning Fraught with Ought? *Pacific Philosophical Quarterly*, 90, 535–555.

Whiting, D. (2013). What is the Normativity of Meaning? *Inquiry*, Online first, DOI**:**10.1080/0020174X.2013.852132.

Wikforss, Å. (2001). Semantic Normativity. *Philosophical Studies*, 102, 203–226.

Wittgenstein, L. ([1953] 1958). *Philosophical Investigations*. G. E. M. Anscombe (Ed.). New Jersey: Prentice Hall.

Wright, C. (1980). *Wittgenstein on the Foundations of Mathematics*. London: Duckworth.

Paper 6

MOORE'S PARADOX AND EPISTEMIC NORMS

Patrizio LO PRESTI

ABSTRACT: Why does it strike us as absurd to believe that it is raining and that one doesn't believe that it is raining? Some argue that it strikes us as absurd because belief is normative. The beliefs that it is raining and that one doesn't believe that it is are, it is suggested, self-falsifying. But, so it is argued, it is essential to belief that beliefs *ought* not, among other things, be self-falsifying. That is why the beliefs strike us as absurd. I argue that while the absurdity may consist in and be explained by self-falsification, we have no reasons to accept the further claim that self-falsifying beliefs are absurd because violating norms.

KEYWORDS: Moore's paradox, epistemic norms, normative explanation, absurdity

1. Moorean Absurdity

G.E. Moore[1] said that there's something 'absurd' with asserting, "It is raining but I don't believe that it is raining." Moore also found believing "He has gone out, but he hasn't" absurd. He found it paradoxical that the absurdity persists despite the possible truth of the proposition asserted or believed.[2] There are circumstances in which it is true both that it is raining and that I do not believe that it is raining. However, it appears absurd to assert, or believe, that it is raining and that I don't believe it. That, in a nutshell, is Moore's paradox.

Moore's paradox displays two faces: a linguistic and a psychological.[3] The linguistic paradox is that it might be true both that it is raining and that I don't believe it although it would be strange of me to *assert* both. The psychological paradox is that it might be true both that it is raining and that I don't believe that

[1] G.E. Moore, "A Reply to My Critics," in *The Philosophy of G. E. Moore*, ed. P.A. Schlipp (New York: Tudor Publishing, 1942), 533-677. See also G.E. Moore, "Russell's Theory of Descriptions," in *The Philosophy of Bertrand Russell*, ed. P.A. Schlipp (New York: Tudor Publishing, 1944), 175-225.

[2] Thomas Baldwin, *G. E. Moore: Selected Writings* (London: Routledge, 1993). This point has also been made in D.M. Rosenthal, "Self-Knowledge and Moore's Paradox," *Philosophical Studies* 77 (1995): 195-209.

[3] Jordi Fernández, "Self-Knowledge, Rationality and Moore's Paradox," *Philosophy and Phenomenological Research* 71, 3 (2005): 533-556; Sydney Shoemaker, "Moore's Paradox and Self-Knowledge," *Philosophical Studies* 77 (1995): 211-228.

it is, although it would be strange for me to *believe* both.[4] I will focus on the psychological version of the paradox.

Both faces of Moore's paradox display two profiles. We may distinguish between believing that

(1) p & I don't believe that p,

and,

(2) p & I believe that not-p.[5]

If you believe (1), you believe that p and that you don't believe that p. If you believe (2), you believe that p and that you believe that not-p. Both (1) and (2) thus involves a first-order belief, that is the first conjunct, and a second-order belief about the first-order belief, that is the second conjunct. In (1) the second-order belief is the belief that you lack belief in the first conjunct. Let us call this the *omissive* version of the paradox. In (2), in contrast, the second-order belief is the belief that you believe the negation of the first-order belief. Let us call this the *commissive* version of the paradox.[6]

I will assume, what is widely agreed, that belief distributes over conjunction.[7] According to the distribution principle, if I believe that it is raining and that water consists of H_2O, I believe that it is raining and I believe that water consists of H_2O.

Distribution Principle: If I believe (p & q), then I believe that p and I believe that q.

From the Distribution Principle we may infer that if I believe the omissive (1), then

(3) I believe that p & I believe that I don't believe that p.

From the Distribution Principle we may also infer that if I believe the commissive (2), then

[4] Rodrigo Borges, "How to Moore a Gettier: Notes on the Dark Side of Knowledge," *Logos & Episteme* V, 2 (2014): 133-140.

[5] Mitchell S. Green and John N. Williams, "Moore's Paradox, Truth and Accuracy," *Acta Analytica* 26 (2011): 243-255; John N. Williams, "Moore's Paradox, Evan's Principle and Self-Knowledge," *Analysis* 64, 4 (2004): 348-353; John N. Williams, "Moore's Paradox and the Priority of Belief Thesis," *Philosophical Studies* 165 (2013): 1117-1138.

[6] Green and Williams, "Moore's Paradox, Truth and Accuracy."

[7] John N. Williams, "Moore's Paradox in Belief and Desire," *Acta Analytica* 29 (2014): 1-23; John N. Williams, "Wittgenstein, Moorean Absurdity and its Disappearance from Speech," *Synthese* 149, 1 (2006): 225-254.

(4) I believe that p & I believe that I believe that not-p.

Both (3) and (4) conserve the initial intuition that Moore-paradoxical beliefs are not first-order contradictions. In (3), since I have a first-order belief that p but not a second-order belief that I believe that p, the second-order belief that I believe that I don't believe that p does not contradict my first-order belief that p. Similarly, in (4), since I have a first-order belief that p but no second-order belief that I believe that p, the second-order belief that I believe that not-p does not contradict my first-order belief that p.

Hence, contradiction arises only by introduction of commutability of a double-belief principle, also known as the principle of Introspective Infallibility,[8] namely:

Introspective Infallibility: If I believe that I (do not) believe that (not-) p then I (do not) believe that (not-) p.

By the principle of Introspective Infallibility we may infer that (3) is self-contradictory. The reason for this is that, under introspective infallibility, the second conjunct's second-order belief (the belief that I don't believe that p) collapses into a first-order omission of belief that p. But this, given the distribution principle, contradicts the first conjunct's first-order belief that p.

We may also infer that (4) is self-falsifying. The reason for this is that, under introspective infallibility, the second conjunct's second-order belief (the belief that I believe that not-p) collapses into a first-order belief that not-p. But this, given the distribution principle, falsifies the first conjunct's first-order belief that p.

The absurdity of Moore-paradoxical beliefs is now clear. The *contents* of the relevant beliefs have unproblematic truth-conditions. But *believing* that one has the beliefs is problematic. If one believes that one's Moore-paradoxical beliefs are true, then, by the Distribution Principle and the principle of Introspective Infallibility, either one has self-contradictory or self-falsifying beliefs. Hence we may conclude with Green and Williams that,

(6) The absurdity of Moore-paradoxical beliefs consists in either self-contradiction or self-falsification.

Note that what constitutes absurdity is not that the beliefs are necessarily false. (1) and (2), for all that (6) says, may be true. It is just that *if* one believes that one's belief in (1) or (2) is true, one's beliefs are either self-contradictory or self-falsifying.

[8] Williams, "Moore's Paradox in Belief and Desire," 5.

Perhaps one disagrees with (6) on the basis that belief distribution or introspective infallibility is false. I will not attempt such an attack here. I will be concerned with suggested *explanations* of the absurdity, rather than with questioning the suggested *constitution* claims. The specific explanation I will argue against is the normative explanation that one *ought* or *may* not have the relevant beliefs. To that end I will grant proponents of such an explanation the premises needed to arrive at (6) – namely both the Distribution Principle and the principle of Introspective Infallibility. Let us grant, then, that the absurdity of Moore-paradoxical beliefs *consists in* self-contradiction or self-falsification, pending whether it is the omissive or commissive form that is at issue.

2. Beliefs and Norms

It has been suggested that what explains the absurdity of Moore-paradoxical belief is epistemic norms. Epistemic norms impinge oughts on doxastic states in general. There are many proposals about precisely what is normative about doxastic states.[9] To understand what about epistemic normativity could possibly account for Moorean absurdity we have first to disambiguate the sweeping claim that belief is normative. That is what I turn to in this section.

Norms are usually supposed to be *imperatives*. For instance, the norm not to cheat has the imperative form: you *ought not* cheat. Some norms may be *conditional* imperatives. For instance, there may be a fairness norm to share with those who have less. This norm has the conditional imperative form: *if* S has less than you, *then* you ought to share with S. The deontic force of the imperative characteristic of norms is not necessarily obligatory though.[10] Instead of impinging *oughts*, a norm may have the force of a *may*; instead of having obligatory deontic force norms may have permissible deontic force.[11] The fairness norm with obligatory deontic force would make it normatively incorrect to not share with those who have less. In contrast, if the same norm had permissibility-force it would not be normatively incorrect to not share with those who have less, since in that case the norm states that you *may* share with those who have less, not that you *ought* to. Not sharing in that case is to not do what you're permitted to.

Epistemic norms likewise impinge imperatives on doxastic states. The deontic force of epistemic norms may be conditional or not, and they may apply to

[9] Clayton Littlejohn, "Are Epistemic Reasons Ever Reasons to Promote?" *Logos & Episteme* IV, 3 (2013): 353-360.

[10] Pascal Engel, "Sosa on the Normativity of Belief," *Philosophical Studies* 166 (2013): 617-624.

[11] Clayton Littlejohn, "Moore's Paradox and Epistemic Norms," *Australasian Journal of Philosophy* 88, 1 (2010): 79-100.

doxastic states as obligations or permissions.[12] But apart from their formal deontic properties, there's also an important difference between the *contents* of epistemic norms. The content of epistemic norms depends on what aspect of doxastic states they are supposed to govern; truth-value, evidential support, justification, etc. These distinctions open up a logical space for a fauna of epistemic norms. The first to be considered among epistemic norms is the truth norm in obligatory form:[13] namely,

> *Obligatory truth norm*: You ought to believe that p only if p.

This norm obliges one to believe only what is true – even if not *all* truths, since it does not have the form 'if p you ought to believe that p.'[14] The obligatory truth norm can be translated into permissive form,[15] thus:

> *Permissive truth norm*: You may believe that p only if p.

The difference between the obligatory and permissive force of these norms may be brought out by substitution of the positive obligatory with obligatory negative form. In that case the obligatory imperative 'ought' translates into the conditional imperative 'ought *not* believe that p *unless* p.' This negative form is imperatively equivalent in force to the positive permissive. According to the latter, you are permitted to believe that p only if p, which is equivalent to being obliged not to believe that p unless p.

A second epistemic norm to consider is the evidence norm,[16] namely:

> *Obligatory evidence norm*: You ought to believe that p only if you have sufficient evidence that p.

The 'sufficient evidence' criterion may be cashed out in a variety of manners depending on one's analysis of 'evidence.'[17] Suppose I believe that it is raining in Reykjavik. One way for my belief to be in accord with the obligatory evidence norm is if I observe the rain myself, if I hear meteorological reports that it is raining in Reykjavik, etc. We may accept that some state or proposition *e*

[12] Anthony Booth and Rik Peels, "Epistemic Justification, Rights, and Permissibility," *Logos & Episteme* III, 3 (2012): 405-411.
[13] Paul A. Boghossian, "The Normativity of Content," *Philosophical Issues* 13 (2003): 31-45.
[14] Nishi Shaw and J. David Velleman, "Doxastic Deliberation," *The Philosophical Review* 114, 4 (2005): 497-534; Pascal Engel, "Belief and Normativity," *Disputatio* 2, 23 (2007): 179-203.
[15] Littlejohn, "Moore's Paradox and Epistemic Norms."
[16] Timothy Williamson, *Knowledge and its Limits* (Oxford: Oxford University Press, 2000); cf. Claudio de Almeida, "What Moore's Paradox Is About," *Philosophy and Phenomenological Research* 62, 1 (2001): 33-58.
[17] Engel, "Belief and Normativity," 185.

qualifies as evidence that p only if it raises the probability that p above some threshold integer, or only if it raises the probability of p above the probability of p in the absence of *e*.[18] Either way the idea is that you ought to believe that p only if the probability of p given *e* meets the relevant qualifier for *e*. The corresponding permissive force of the evidence norm is,

> *Permissive evidence norm*: You may believe that p only if you have sufficient evidence that p.

This norm differs from the former obligatory in that if you do *not* believe that p given *e* you're not normatively incorrect, since in this case you're simply not exerting permission. In the former obligatory form this would be incorrect though. For in that case you do not just not utilize permission but violate an obligation.[19]

It may, thirdly, be suggested that *knowledge* is an epistemic norm for doxastic states.[20] The knowledge norm with obligatory deontic force reads:

> *Obligatory knowledge norm*: You ought to believe that p only if you know that p.

The imperatival force of this norm is that your belief that p is as it ought to be just in case you know that p is true. The knowledge-norm thus differs from the truth-norm in there being circumstances in which your belief that p is in accord with the latter but in violation of the former. There are circumstances in which your belief that p is true but you don't know it.[21] Translating the knowledge norm into its permissive counterpart, we get:

> *Permissive knowledge norm*: You may believe that p only if you know that p.

It should be clear by now in what the difference between the obligatory and permissive force of the relevant norm consists. In the former obligatory you are wrong in not believing that p if you know that p whereas, in the latter permissive, you are not wrong if you don't believe that p when you know that p since it says that you *may* believe that p only if you know that p. The permissive knowledge norm and the permissive truth norm differ in a similar manner to how their

[18] Franck Lihoreau, "Are Reasons Evidence of Oughts?" *Logos & Episteme* III, 1 (2012): 153-160.
[19] Conor McHugh, "Beliefs and Aims," *Philosophical Studies* 160 (2012): 425-439.
[20] Littlejohn, "Are Epistemic Reasons Ever Reasons to Promote?"; Declan Smithies, "The Normative Role of Knowledge," *Noûs* 46, 2 (2012): 265-288; Michael Huemer, "Moore's Paradox and the Norm of Belief," in *Themes from G. E. Moore: New Essays in Epistemology and Ethics*, eds. Susana Nuccetelli and Gary Seay (New York: Oxford University Press, 2007), 142-158.
[21] David Owens, *Reasons Without Freedom* (London: Routledge, 2000); David Owens, "Does Belief Have an Aim?" *Philosophical Studies* 115 (2003): 283-305.

obligatory versions differ. That is, there are circumstances in which a belief that p is permitted courtesy p being true whereas, if p is not known, believing p in those circumstances violates the permissive truth norm.[22]

Note that the above kinds of norms may support each other in various ways.[23] Endorsement of either version of the knowledge norm naturally supports endorsement of the conjunction of the corresponding version of the truth and evidence norms. The reason for that is that if you accept that it is correct to believe that p only if one knows that p, then, on most accounts of knowledge, p must be true and the belief that p enjoy some kind of support.[24] This norm-conglomeration is not necessary though. You may endorse either version of the evidence norm, for instance, yet deny both versions of the truth norm on the grounds that given accord with the former your belief is permitted even if false.[25] Then again, you may argue that there's no *absolute* norm of belief but that beliefs may be correct or incorrect in many different respects simultaneously.[26] In some circumstances the normative correctness of doxastic states may be adjudicated by their truth-value while, in others, it may be adjudicated by evidential support.

We have now distinguished epistemic norms according to their *contents* – whether the aspect of doxastic states that the norms are about is truth-value, evidential support, or knowledge – and according to *deontic force* – whether the norms take obligatory and permissive form. We have also considered the possibility of combining these in various respects. But epistemic norms may be distinguished along a further, third axis, namely, according to in what relation doxastic states are supposed to stand to the different imperatives. Irrespective of a norm's content and force we may ask how the norm applies to doxastic states to begin with. Suppose, for instance, that I believe that water has the chemical composition CH_4. Then you tell me that I ought not have that belief because it is false. I might then wonder what the nature of the purported relation between my belief and the norm is. There are basically two alternative understandings of how imperatives attach to doxastic states.

One proposal is that the nature of the relation between beliefs and norms is conceptual.[27] On this account, it is analytically true that a belief is correct only if it is in accord with the relevant norm.

---

[22] Williamson, *Knowledge and its Limits*.

[23] Littlejohn, "Moore's Paradox and Epistemic Norms."

[24] Williamson, *Knowledge and its Limits*.

[25] Boghossian, "The Normativity of Content."

[26] José L. Zalabardo, "Why Believe the Truth? Shah and Velleman on the Aim of Belief," *Philosophical Explorations* 13, 1 (2010): 1-21.

[27] Engel, "Sosa on the Normativity of Belief," 621.

> *Conceptual claim*: The concept of belief is such that the belief-norm applies to all beliefs.

Suppose, e.g., that the norm under consideration is the truth norm in obligatory form. According to the conceptual claim[28] the norm would read:

> *Conceptual obligatory truth norm*: The concept of belief is such that, for all beliefs, you ought to believe that p only if p.

My belief that water is CH_4 would then be incorrect according to our understanding of the concept 'belief.' We cannot understand something as a belief, the suggestion is, without understanding it as something one is obliged to if true, thus as incorrect if false. By believing that water is CH_4 I violate an obligation to believe only truths – an obligation attached to belief by definition. Consider in contrast the truth norm with permissive force.[29] From the conceptual claim we then get,

> *Conceptual permissive truth norm*: The concept of belief is such that, for all beliefs, you may believe that p only if p.

In this case my belief that water is CH_4 is, again, incorrect according to how we conceptualize belief, because I am not permitted to that belief given that water is not CH_4. However, *were* water CH_4 but I did not believe it, the omission of belief would not be incorrect, since I would then merely have not utilized a permission to believe.

The evidence and knowledge norms are easily translatable into the conceptual claim. All we have to do is to substitute them for 'the belief-norm' in the conceptual claim. I will not waste space making them explicit here. All that is required is to insert 'the concept of belief is such that…' before the imperative 'ought' or 'may' in the relevant norm above. This would yield the norm that, for instance, the concept of belief is such that you may believe that water is CH_4 only if there is some proposition *e* such that the probability that water is CH_4 given *e* is higher than water not being CH_4.

The other answer to our inquiry into the nature of the relation between alleged epistemic norms and doxastic states is that the relation is metaphysical. It is claimed that the *nature* of the psychological state that is belief is such that it is normatively regulated.[30] This metaphysical connection is often spelled out in term

[28] Shah and Velleman, "Doxastic Deliberation," 252; David Velleman, *The Possibility of Practical Reason* (Oxford: Clarendon Press, 2000), 277-278.
[29] Littlejohn, "Moore's Paradox and Epistemic Norms."
[30] Ralph Wedgwood, "The Aim of Belief," *Philosophical Perspectives* 16 (2002): 267-297.

of the 'aim of belief.'[31] The aim of cognitive mechanisms responsible for belief production are such that, as the familiar slogan has it, their aim is the production of a state whose representational content fits the world.[32] If the produced state's contents don't fit the world it is incorrect. Another slogan that quite pinpoints the idea is of normative principles 'built into' our cognitive apparatuses.[33] We may formulate the relevant connection thus:

> *Metaphysical claim*: The nature of belief is such that the belief-norm applies to all beliefs.

The procedure of disambiguation of various contents and force of metaphysical belief-norms should be clear by now. Substituting the truth-, evidence- or knowledge norm in either obligatory or permissive form for 'the belief-norm' in the metaphysical claim yield the corresponding specification. For example, introducing the truth norm with obligatory deontic force gives,

> *Metaphysical obligatory truth norm*: The nature of belief is such that, for all beliefs, you ought to believe that p only if p,

and so on for the other norms and deontic forces. To avoid tedious repetitions I'll avoid spelling out their exact formulations here. If necessary we may do so at any point in the argument.

The difference between the conceptual and metaphysical construal of the relation between epistemic norms and doxastic states is this. The conceptual claim entails that possession of the concept of belief is sufficient for a subject to recognize that, were his belief that p to violate the relevant norm, then his belief would be normatively incorrect.[34] What explains incorrectness in this case is the norm analytic to the concept of belief. On the metaphysical construal, in contrast, insofar one has, say, a representation-dedicated cognitive module with the aim of supplying truth-valued representations,[35] then satisfaction of that aim suffices for the output cognitive states to be in accord with the relevant norm. Here it is the

---

[31] Benjamin W. Jarvis, "Norms of Intentionality: Norms that don't Guide," *Philosophical Studies* 157 (2012): 1-25.

[32] Bernard Williams, *Problems of the Self* (New York: Cambridge University Press, 1973); Asbjorn Steglich-Petersen, "No Norm Needed: On the Aim of Belief," *The Philosophical Quarterly* 56, 225 (2006): 499-516; Asbjorn Steglich-Petersen, "Weighing the Aim of Belief," *Philosophical Studies* 145 (2009): 395-405.

[33] Ralph Wedgwood, "The A Priori Rules of Rationality," *Philosophy and Phenomenological Research* 59, 1 (1999): 113-131.

[34] Pascal Engel, "Is Truth a Norm?" in *Interpreting Davidson*, eds. Petr Kotátko, Peter Pagin, and Gabriel Segal (Stanford: CSLI Press, 2001), 37-51.

[35] Wedgwood, "The A Priori Rules of Rationality," 130.

nature of the state that determines its correctness conditions, or vice versa, depending on the order of metaphysical determination alleged to obtain between epistemic norms and doxastic states.[36] It may be suggested that norms determine the nature of the state, or, the other way round, that the nature of the state determines what norms apply. Either way, when it comes to the analyticity of norms of belief suggested by the *conceptual claim*, the *nature* of the state as such is secondary to the application of the norm, while it is the other way round for the *metaphysical claim*. According to the latter, whatever *definition* we use to distinguish beliefs from other psychological states beliefs are different ultimately with reference to the 'aim' or 'goal' that govern their production.

To conclude this section, we find that the claim that belief is normative admits of a multitude of specifications. Normativity claims, unless properly disambiguated, are quite sweeping. I have tried to provide some specifications here. According to the specifications provided, there are three kinds of norms, each with an obligatory and a permissive form that might be understood as conceptually or metaphysically related to doxastic states. This basically yields twelve versions of belief-normativism (if we abstract from combinations of kinds of norms, such as the knowledge- and truth-norms). We're now in a position to home in on and criticise various claims that the reason why Moore-paradoxical beliefs are absurd is that they violate epistemic norms.

3. First Attempted Normative Explanation of Absurdity

Green and Williams[37] suggest that the absurdity of Moore-paradoxical belief consists in severe violations of belief-norms: "Do not form – or continue to have – a specific belief that you can reasonably expect to be false" and "Do not form – or continue to have – a specific belief that you can be reasonably expected to see is self-falsifying."[38] These are norms that any "epistemically rational" believer "certainly would endorse."[39] Epistemic rationality is to be understood as "that property of one's acquiring or continuing to have it [the belief] that turns it, if true and not Gettierized, into knowledge."[40]

Commissive Moore-paradoxical beliefs are suggested to violate the norm not to form or continue to have self-falsifying beliefs. Therefore, this account has

[36] Kathrin Glüer and Åsa Wikforss, "Against Content Normativity," *Mind* 118, 469 (2009): 31-70.

[37] Green and Williams, "Moore's Paradox, Truth and Accuracy," 249.

[38] Williams, "Moore's Paradox and Priority of Belief Thesis," 14.

[39] Williams, "Moore's Paradox and Priority of Belief Thesis," 14.

[40] Williams, "Moore's Paradox in Belief and Desire," 2.

it, Moore-paradoxical beliefs are absurd. Moore-paradoxical beliefs are not necessarily irrational, though. One will not always be "in a position to see" that one's beliefs are in violation of the relevant norms.[41] So absurdity is a violation of belief-norms but it isn't surface-level self-falsification. This seems right. It preserves the conclusion arrived at in the first section, that Moorean absurdity is a property of conjugated second- and first-order beliefs that falsify or contradict each other under distribution and introspective infallibility.[42] I agree with Green and Williams up to (6). We agree that if I form or continue to have the commissive Moore-paradoxical belief,

(2) p & I believe that not-p,

then, by introducing the Distribution Principle,

(DP 2) I believe that p & I believe that I believe that not-p,

which, given the principle of Introspective Infallibility, yields:

(7) I believe that p & I believe that not-p.[43]

The conjuncts of the belief falsify one another. To arrive at this conclusion we've granted Green and Williams the auxiliary principles of distribution and infallibility they need. In other words, we are in agreement that what *constitutes* absurdity is that the beliefs are self-falsifying. But Green and Williams makes a further claim. The further claim is that what *explains* the absurdity is violations of belief-norms.[44] Here I find reason to disagree.

The relevant norm is that one ought not form or continue to have beliefs that are self-falsifying.[45] Given that commissive Moore-paradoxical beliefs are self-falsifying they violate the relevant norm. That is why, Green and Williams argue, the beliefs are absurd. Green and Williams's normative account should be rejected for the reason that one might accept that the beliefs are absurd because self-falsifying while rejecting that self-falsifying beliefs are norm-violations. We may agree that what *constitutes* the absurdity of commissive Moore-paradoxical beliefs is that their contents are in tension, granted the agreed upon premises. And so we may answer the question *why* a commissive Moore-paradoxical belief is absurd by pointing out that forming or continuing to have it is to form or continue to have a

[41] Green and Williams, "Moore's Paradox, Truth and Accuracy," 250.

[42] Williams, "Moore's Paradox in Belief and Desire," 5.

[43] Williams, "Moore's Paradox in Belief and Desire," 6.

[44] Green and Williams, "Moore's Paradox, Truth and Accuracy," 249-250; Williams, "Moore's Paradox and the Priority of Belief Thesis," 15-16.

[45] Williams, "Moore's Paradox in Belief and Desire," 6-7.

belief whose conjuncts falsify each other. The absurdity is then *explained* by the fact that *believing* that p and that one believes that not-p, collapses, given the Distribution Principle and the principle of Introspective Infallibility, on which we agree for the sake of argument, into a self-falsifying belief. But that there is an *additional* reason why the beliefs are absurd, namely because an epistemic *norm* not to form or continue to have the relevant beliefs is violated, finds no support in the argument.

To illustrate, consider the beliefs that, say, it is raining and that it is not raining. Suppose I form or continue to have both. I then have self-falsifying beliefs. If I believe one then the other must be false. Now, my *reasons* for forming or continuing to have both beliefs or, indeed, my reasons for not forming both or for abandoning either, might be a range of reasons none of which necessarily is the reason that I *ought* or ought *not* to form or continue to have both. What constitutes the absurdity appears to be that the beliefs are self-falsifying. That is all well and good. But in order for it to be true that what *explains* the absurdity is a violation of epistemic norms it is necessary that at least part of what does the explanatory work is my having a reason that I *ought* or ought *not* to form or continue to have the beliefs. I do not violate or conform to a norm if, by chance, I *happen* to be wrong or right. It should rather be the case that, if we're interested in *normative* explanation, I form or continue to have the beliefs in question because I recognize that I ought *not* or *ought* to form or continue to have them.

As far as Green and Williams's argument is concerned, and I see no reason to disagree, nothing suggests that part of anyone's reasons for forming or continuing to have Moore-paradoxical beliefs is that they *ought* or ought *not* to. Admittedly, Green and Williams suggest that it is only if one recognizes that one's beliefs would be self-falsifying that they are absurd. But, surely, one might recognize that one has absurd beliefs in the sense of their being self-falsifying without it *also* being the case that one has the beliefs even partly for the reason that one ought or ought not to. Hence, Green and Williams might be entirely right that the absurdity of Moore-paradoxical belief consists in self-falsification, yet not thereby having provided any reason for accepting that the absurdity is explained by violations of epistemic norms.

It may be objected, by those of normativist persuasion, that belief, the psychological state as such, still 'aims at truth,' or is 'directed to fit' the world.[46] And, in that sense of 'normative,' beliefs that fail to meet this aim or that don't fit the world, as is the case of Moore-paradoxical beliefs *if* believed to be true by the

[46] Cf. Daniel Whiting, "Does Belief Aim (Only) at the Truth?" *Pacific Philosophical Quarterly* 93, 2 (2012): 279-300.

believer, would be wrong no matter for what reasons the believer form or continue to have them. More generally, the point may be the metaphysical normative claim that if a psychological state does not have the relevant aim, is not governed by the relevant direction of fit, then it is not a belief. And, if it fails to meet its aim, or fails to 'fit,' then it is normatively incorrect no matter what the reasons are for which the believer forms of continues to have the relevant beliefs. To this I respond that we may accept that beliefs necessarily 'aim at truth' or 'aim to fit the world,' and that any epistemically rational believer would accept this.[47] Beliefs that fail to meet this aim would be, let us say, incorrect or wrong. However, if 'incorrect' and 'wrong' in this context is not to be understood in relation to the believer's normative reasons, then 'wrong' and 'incorrect' can be made perfect sense of as *descriptive* terms. Straightforwardly, false beliefs are 'incorrect' precisely because false.[48] It would be untoward to speak of false beliefs that aim at truth, but not necessarily for any subjective normative reason, as incorrect or wrong because they *violate* oughts. A belief as such does not violate anything; it is true or false. Only by recognizing, but going against, a reason can one violate it. A belief, however, does not have reasons for its own formation or maintenance, much less *normative* reasons. Hence, if a belief is true or false it may be correct or incorrect in the descriptive sense. But if it is not for any normative reason that 'correct' and 'incorrect' apply, there seems to me nothing left from which a *normative* understanding of 'correct' and 'incorrect' can derive plausibility. Therefore, this objection fails. Moving to the metaphysical normative claim in defence of a normative explanation of Moore-paradoxical beliefs is to move away from whatever may originally have lent such an explanation support.

In a similar vein of response to the metaphysical move, I may believe that it is raining yet believe that I do not believe this, and perhaps be self-contradictory and 'absurd,' for a number of reasons. But this does not suffice for the *additional* claim that, nor does it seem necessary for the claim that, I have any particular *normative* reason stating that I ought or ought not form or continue to have the beliefs. Therefore, even accepting metaphysical claims about the 'aim' of belief, no normative constituency claim about, or normative explanation of, false beliefs follow. Likewise, the absurdity of self-falsifying beliefs, as we assume that some Moore-paradoxical beliefs are, would still not consist in or necessitate an explanation in terms of norm-violation. At least, insofar we agree with Green and Williams's premises, no normative conclusion follows.

---

[47] Green and Williams, "Moore's Paradox, Truth and Accuracy."

[48] Glüer and Wikforss, "Against Content Normativity," 35-36; Fred Dretske, "Norms, History and the Mental," *Royal Institute of Philosophy Supplement* 49 (2001): 87-104.

Perhaps it will still be objected that Green and Williams's point is that the absurdity of Moore-paradoxical beliefs consists in and is explained by norm-violations if I can be *reasonably expected to recognize* that the beliefs would be self-falsifying.[49] As Williams puts it, "I violate the norm … because I may be reasonably expected to see that my belief is self-falsifying."[50] There are two reasons why the response does not avail the normativity part of Green and Williams's approach.

The first reason is that Moore-paradoxical beliefs appear no *less* absurd merely because one does not to recognize that they would be, say, self-falsifying, and thereby in violation of alleged norms. My belief, e.g., that it is raining and that I believe that it is not raining, bears the hallmark of absurdity because, we are assuming, it is self-falsifying. It would be absurd even if I do not *also* recognize that it is self-falsifying and even if I do not *also* recognize that the belief would violate some alleged norms of belief. Similarly, it appears no *less* 'correct' to reject that it is not raining if I believe that it is raining than it would be 'correct' to do so *and* do it because I recognize that one *ought* to. The beliefs are absurd or not quite irrespective of one *also* recognizing that they violate or conform to norms.[51] Hence, the suggestion that it is only if I recognize that my Moore-paradoxical beliefs would be in violation of epistemic norms that my Moore-paradoxical beliefs are absurd does not avail Green and Williams's account.

The other reason for rejecting the present response is that a vicious regress ensues if the absurdity of Moore-paradoxical beliefs consists in forming or continuing to have them despite recognizing that in forming or continuing to have them one violates belief-norms. To 'recognize' beliefs as violating norms requires, minimally, believing that they would violate the relevant norms. If this is *not* required, then it cannot be *because* of norm-*violations* that one's beliefs are absurd, because a necessary means to *violate* is to believe that one ought (not) form some belief, yet, despite this, form (or not form) it.

To demonstrate how the regress will inevitably be engendered if we accept the normative part of Green and Williams's account, suppose that we grant their point that it is only by recognizing the normative incorrectness of one's beliefs, yet continuing to have them, that the beliefs are absurd. From this we may infer that beliefs are absurd only if one has a second-order belief that the beliefs are incorrect. That is, unless one is in a position to recognize, i.e., minimally, believe,

[49] Williams, "Moore's Paradox in Belief and Desire," 7.

[50] Williams, "Moore's Paradox in Belief and Desire."

[51] Kathrin Glüer and Åsa Wikforss, "Aiming at Truth: On the Role of Belief," *Teorema* 32, 3 (2013): 137-162.

that one's beliefs are incorrect then they are not absurd. But now the necessary *second-order belief* about the normative incorrectness of one's first-order beliefs, qua itself a belief, will, on the normative proposal, be subject to the relevant epistemic norms. The second-order belief about the incorrectness of any first-order belief may itself be absurd, if it violates epistemic norms and I am in a position to recognize that this is the case. (In fact, if my second-order belief so much as *could* give rise to absurdity, *absent* recognition that it violates some alleged norm, then Green and Williams's account will be falsified. For in that case there is absurdity that does not consists in or is explained by norm-violations. On the other hand, if the second-order belief is *not* susceptible to epistemic norms just like the first-order beliefs, then the normative account will also be falsified. For then we have beliefs that may be false or self-falsifying yet not violate norms.) Suppose now that I form the necessary second-order belief about the normative incorrectness of my Moore-paradoxical beliefs and I recognize that my Moore-paradoxical beliefs would be in violation of epistemic norms. The obvious question then is: Is my second-order belief that my Moore-paradoxical belief is normatively incorrect itself normatively correct or incorrect? If we accept Green and Williams's normative account, then we can explain the absurdity or lack of absurdity of my second-order belief only by settling whether I recognize, i.e., minimally, believe, that it violates (or not) the relevant epistemic norms. I now form the necessary third-order belief about the normative correctness or incorrectness of my second-order belief that my Moore-paradoxical beliefs are normatively incorrect…[52] Again, assuming that it is *possible* that higher-order beliefs are false or self-falsifying in relation to the lower-order beliefs that they are about, we again face the dilemma of settling whether the higher-order belief is absurd or not. If *it*, the third-order belief, cannot be absurd or not, then it is not the case that belief is normative. In that case, the prospects for providing a *normative* account of Moorean absurdity dims significantly. But if the higher-order beliefs can themselves be absurd for the normative reasons defended by Greens and Williams, then they would be absurd because I recognize, i.e., minimally, believe, that they violate some epistemic norm. In that case the regress pushes us towards absurdity for as long as we maintain that Moorean absurdity consists in or is explained by epistemic norms in the sense advocated by Green and Williams.

[52] Note that what I refer to as a second-order belief in this argument is actually a third-order belief, and the third-order belief a fourth-order belief. The reason for this is that a Moore-paradoxical belief itself embodies a second-order belief about a first-order belief. So any belief about Moore-paradoxical beliefs will begin at the third-order.

The absurd consequence of the normative part of Green and Williams's account is, then, that in order for Moore-paradoxical beliefs to *be* absurd the believer must form ever higher-order beliefs. The only way out of this dilemma is to recognize that belief, including Moore-paradoxical beliefs, may be absurd for no normative reason but because they are, say, self-falsifying or self-contradictory. This is what Green and Williams's argument shows.

4. Second Attempted Normative Explanation

In this section we find reasons to reject another suggestion that Moorean absurdity consists in and is explained by violations of epistemic norms. The suggestion is due to Pascal Engel. He writes,

> The reason why they [Moore-paradoxical beliefs] are paradoxical and the reason why we hesitate to attribute to the agent both the belief that P and the belief that not P is that when someone has a belief that P, he thereby has the belief that P is *true*. If he comes to believe (consciously, at the same time) that his belief that P is false, then either he does not have either one belief, or he is *not really* … in a state of belief. So even someone who, for any reason, is not moved by an interest for truth, or who rejects the idea that it can be a goal for his beliefs, has to recognize that truth is what his beliefs are aiming at, *in virtue of their being beliefs*.[53]

To be fair, Engel's general aim in this context is not to explain the absurdity of Moore-paradoxical beliefs. Even so, the passage is illuminating. In a recent paper Engel adds that, that beliefs aim at truth "is not true in the descriptive or causal sense … It has to be true in the sense of conceptual necessity, or of normative necessity."[54] There are three points worthy to highlight in the quoted passage.

First, note that Engel says that we hesitate to attribute Moore-paradoxical beliefs because it involves attributing "both the belief that p and the belief that not-p." Engel seems to misunderstand Moore-paradoxical beliefs. To begin with, not all Moore-paradoxical beliefs have this form. In some cases, namely in the omissive version of the paradox, the belief is (1) "p but I don't believe that p." Engel does not mention this. His next point, that if one believes that p then one believes that p is true, will thus not apply to Moorean absurdity in general. On the other hand, the commissive form of the paradox that Engel mentions, in particular, has the form of (2) "p but I believe that not-p." What you should attribute to me if I have *this* belief is not the first-order beliefs that p and that not-

[53] Engel, "Is Truth a Norm?" 49 (emphases added).

[54] Engel, "Sosa on the Normativity of Belief," 621; Shah and Velleman, "Doxastic Deliberation," 525; Shah, "How Truth Governs Belief," *The Philosophical Review* 112, 4 (2003): 447-482.

p, but the first-order belief that p and the *second-order* belief that I believe that not-p. Otherwise the Distribution Principle yields surface-level contradiction. And that, it is clear, is not what Moore's paradox is about.[55] Perhaps Engel implicitly assumes that the principle of Introspective Infallibility is correct. We have, for the sake of argument, granted that principle. Even so, one would require some further support of it if it were to carry the weight it does here. However, let us grant again, for the sake of consistency, that the principle of Introspective Infallibility is correct.

Engel's approach then faces a second dilemma. He claims that if I come to believe that one of my Moore-paradoxical beliefs is false, then either I' don't *really* have both or I'm not *really* in a state of belief with regard to one of them. In a sense then, I cannot really have both beliefs. And this is true "in the sense of conceptual necessity, or of normative necessity." That is, since there are cases in which we in fact fail to 'hit at' truth when forming beliefs, it is not the case that we *do* believe only truths, but that we *ought* to believe only truths.[56] Engel's proposal, then, is this. My belief that

> (2) p and I believe that not-p

is absurd because

> *Conceptual Truth Norm*: The concept of belief is such that, for all beliefs, you ought to believe that p only if p.

Furthermore, we saw that Engel deploys what we might call the thesis of Normative Resistance:

> *Normative Resistance*: If you believe that your belief that 'p and I believe that not-p' violates the conceptual truth norm, then either you do not *really* believe one of the conjuncts or you are not really in a state of *belief* with regard to one of them.

The thesis of Normative Resistance is problematic. If we accept it, then Moore-paradoxical beliefs are impossible. Here's why. According to the thesis, if I believe that p and that I believe that not-p, then either I *cannot* believe both conjuncts, or I'm not in a state of *belief*. In that case I cannot really believe that p and that I believe that not-p. Now, if I cannot believe this then I cannot really have the Moore-paradoxical belief. But what is to be accounted for is precisely the absurdity of beliefs of the form (2) "p, but I believe that not-p." Supposing that Normative Resistance is correct, in conjunction with the conceptual truth norm,

[55] de Almeida, "What Moore's Paradox is About."
[56] Engel, "Belief and Normativity."

makes this belief impossible. This, in turn, is problematic for two reasons. On the one hand, we have established that it is perfectly possible for it to be the case that p and that I believe that not-p.[57] Indeed, and this is the other problem, the situation had better *not* be impossible. For if it were, then the explanation would explain nothing. The explanans – the conceptual truth norm and the thesis of Normative Resistance – render the explanandum – the belief that p and that I believe that not-p – impossible. But then the explanation is itself contradictory. It starts out by having us imagine an instance of Moore-paradoxical beliefs. It then attempts to explain what is paradoxical in terms of norms that apply to beliefs by conceptual necessity. But it thereby renders the beliefs in question impossible by conceptual necessity. And so the approach debouches in the claim that the reason why we hesitate to attribute Moore-paradoxical beliefs is that Moore-paradoxical beliefs are impossible. Hence Engel effectively ends up empty-handed; there's nothing to explain, much less anything meriting normative explanation.

However, let us grant Engel that, somehow, the explanation can nevertheless be made to work. That is, assume that the reason why we hesitate to attribute to an agent a Moore-paradoxical belief is that we would then be attributing beliefs that violate the conceptual truth norm. This leads us to the third dilemma. One hallmark of *norms* is that they tell us what we ought (not) to or may (not) do; i.e., they take the form of imperatives with deontic force. If, as Engel rightly points out,[58] we substitute the 'ought' or 'may' in the imperative for a 'do' or 'will,' then the result is not norms, but descriptions of regularities between facts, evidence and the formation of belief. We can put this point in terms of the requirement that,

> *Normative Difference*: Norms *should make a difference* to the way we form, manage and revise beliefs.

The deontic force embedded in an epistemic imperative should, that is, play a role in our forming, and way of forming, revising and abandoning beliefs. Glüer and Wikforss[59] argue that if a reason for belief fails to satisfy this requirement, then it is redundant to label it a *normative* reason. That is, if no part of one's reason for believing p is the reason that one *ought* to believe that p, then, even if there were a norm for believing that p, the norm makes no difference for what one ends up believing and how one ended up believing it. And if the norm makes no difference, then it is utterly idle and plays no role in an account of (manners of)

[57] Borges, "How to Moore a Gettier," 134.

[58] Engel, "Sosa on the Normativity of Belief."

[59] Kathrin Glüer and Åsa Wikforss, "Against Belief Normativity," in *The Aim of Belief*, ed. Timothy Chan (Oxford: Oxford University Press, 2014), 80-99.

forming, continuing to have or revise belief. Furthermore, the Normative Difference-requirement suggests that one be in a position to choose to believe (or not believe) that p. A normative reason is open to deliberation. If what explains my belief that p is the normative reason that I *ought* to – rather than that I *will* believe it in my circumstances – then I may *violate* the norm. If I *cannot* do so, then the reason for my belief that p cannot be a *normative* reason. At the very least, to say that it was a normative reason that I *could not* violate would be no different from saying that it was, say, a causal reason that I had no influence over. So a norm has to make a difference in the sense that it can figure as my reason to form, maintain or revise a specific belief without it being necessary that I form, maintain or revise the belief accordingly.

Our considerations now make obvious the third dilemma with Engel's normative account. If violations of the conceptual truth norm are cases of *not really believing*, and if this is so by conceptual necessity, then the conceptual truth norm cannot make a difference in my forming, maintaining or revising Moore-paradoxical beliefs. If it is the case that I *cannot really* believe that p and that I believe that not-p, then it does not matter if, in addition, I ought or ought not have these beliefs. Of course I *may* recognize that I ought not have the relevant beliefs. But the *reason why* I don't (indeed never *really*) form them would, on Engel's account, not be that I recognize that I ought not to, but that I cannot, given the conceptual normative truth about belief. Hence, what explains the absurdity of Moore-paradoxical beliefs cannot be some *normative* reason. Perhaps there are such normative reasons against Moore-paradoxical beliefs. But the way Engel has set up the case, no such reasons figure in the explanation of why Moore-paradoxical beliefs are paradoxical.

We may conclude that Engel's normative explanation is problematic for three reasons. First, he doesn't really address Moore's paradox, at least not in its full complexity. Secondly, even if he were to address the paradox he would make it an impossible explanandum since his analysis of 'belief' entails that there cannot really be Moore-paradoxical beliefs. If there cannot be Moore-paradoxical beliefs, then there simply is no (normative) explanation of Moore-paradoxical beliefs. Thirdly, the norm invoked to explain the paradox would be explanatorily idle because if one forms Moore-paradoxical beliefs (given that one could) one would believe incorrectly no matter whether one forms the beliefs for any normative reason. Moreover, if the reason why one does not form Moore-paradoxical beliefs is that one *cannot*, then a norm that one *ought not* ads nothing to why one does not.

5. Conclusions

The two lines of thought examined here, according to which the absurdity of Moore-paradoxical beliefs consists in and is explained by belief-norm violations, are problematic. I have had the courtesy to grant premises that on closer inspection might themselves be problematic. The principle of Introspective Infallibility, for instance, might not be appealing to some. Some might find that Moore's paradox isn't about self-falsifying beliefs at all, as I have granted here. I'm sympathetic to worries that perhaps the paradoxical nature of the beliefs should be understood along other lines. But here I've wished to grant proponents of normative accounts as much as possible in order to refute their case. In being generous, we've found reasons to reject normative accounts. I conclude that an account of the psychological version of Moore's paradox that appeals to epistemic norms is unsatisfactory.[60]

[60] Thanks to participants at the ECAP8 for comments on an earlier version of this paper. Special thanks to Åsa Wikforss for helping me get to the point.