

INTERSUBJECTIVITY AND COMMUNION

ON JOINT ACTION AND SOCIAL ONTOLOGY

Intersubjektivitet og fellesskap

Om fellehandling og sosial ontologi

by

Eirik Moland Fevang



The University of Bergen

Master's thesis in FILO350

Supervisor: Ole Martin Skilleås

Department of Philosophy

Spring 2023

Table of contents

- Abstract 3**
- Introduction 5**
- Chapter 1: Mapping the discussion 7**
 - 1.1 On the contemporary field of joint action 7*
 - 1.2 The question 9*
 - 1.3 Action and intention..... 11*
- Chapter 2: The standard theories..... 13**
 - 2.1 The non-reductive accounts of shared intentions and joint activity..... 13*
 - 2.2 The reductive theory..... 21*
- Chapter 3: The enactive approach..... 29**
 - 3.1 Critical perspectives on the standard theories..... 29*
 - 3.2 Enactive theory on joint attention and language acquisition 39*
 - 3.3 Strengths and weaknesses of enactive theory..... 46*
- Chapter 4: The phenomenological approach 49**
 - 4.1 Introducing classical phenomenology 49*
 - 4.2 The Phenomenology of Intersubjectivity 54*
- Chapter 5: Joint action and communion 62**
 - 5.1 The linguistic dimension 62*
 - 5.2 Joint attention and communion..... 69*
 - 5.3 Expressive theory and joint action..... 73*
 - 5.4 Further considerations..... 76*
- Conclusion..... 80**
- References 82**

I am grateful to Ole Martin Skilleås who has been my supervisor in this project. I also want to express my gratitude to Franz Knappik for helpful suggestions and guidance. To my parents Bjørg-Tilde Svanes Fevang and Jonas Meling Fevang, thank you so much for all encouragement and support. To my sister Anne Berit Fevang, my brother Knut Jonas Sandok Fevang, and my sister-in-law, Mari Sandok Fevang, and to my grandmother, Bjørg Svanes, and my late grandfather, Knut Svanes, I want to express my sincerest thanks for their continued interest and support. To my friends from the Wergeland gang, Kjetil, Peter, Tord, Magnus, Mariell, Hallvard and Annelise, thank you for your friendship and inspiring discussions.. Most of all, I would like to thank my wife, Torill Moland Fevang, whose patience and wisdom has been indispensable during the writing of this thesis.

Abstract

In this thesis, I enquire into what it means to act together, *joint action*, and the ontological conditions for joint action. I discuss the standard non-reductive theories of John Searle, Margaret Gilbert and Raimo Tuomela and the reductive account of Michael Bratman, siding with the latter in arguing that social groups need not activate a special “we-intention” in order to undertake joint action and cooperation. Following this, I discuss the enactivists’ critiques of the standard theories, which underline the role of the body in joint action and social life. Their theories are based on recent studies of joint attention cases involving autistic and non-autistic children and caregivers. While appreciating some insights of enactive theory, certain points are discussed and criticized before I move on to discuss Dermot Moran’s phenomenological theory in which the importance of the *life-world* is essential. Lastly I build on insights from Charles Taylor’s expressive theory in arguing that the irreducible background horizon of the linguistic dimension must be taken into account when discussing joint action, joint attention, and social ontology. Integrating Moran and Taylor’s insights on intersubjectivity, language and the phenomena of *communion* or emotional joint attention, with the planning theory of shared agency provided by Bratman, renders an adequate and comprehensive account of joint action. In Taylor’s theory, language itself constitutes the most fundamental level encompassing all forms of communication from the enactive to the conceptual and intellectual. I conclude by stating that the matrix of language is conversation, and this remains so throughout human history, being fundamental to all joint activity and group intention.

Abstrakt

I denne oppgaven undersøker jeg hva det vil si å handle sammen, dvs. *felleshandling*, og de ontologiske betingelsene for felleshandling. Jeg diskuterer de standard ikke-reduktive teoriene til John Searle, Margaret Gilbert og Raimo Tuomela og den reduktive beretningen til Michael Bratman, og tar side med sistnevnte ved å hevde at sosiale grupper ikke trenger å inneha en spesiell "vi-intensjon" for å gjennomføre felleshandling og samarbeid. Deretter går jeg videre til å diskutere enaktivistenes kritikk av standardteoriene, som legger større vekt på kroppens rolle i felleshandling og sosialt liv. De baserer sine teorier på nyere studier av felles oppmerksomhet mellom autistiske og ikke-autistiske barn og deres omsorgspersoner. Selv om jeg verdsetter noen av innsiktene til enaktiv teori, kritiserer jeg dem og går videre til å diskutere Dermot Morans fenomenologiske teori der betydningen av *livsverdenen* fremheves. Til slutt bygger jeg på innsikt fra Charles Taylors ekspressive teori ved å argumentere for at den språklige dimensjonen er en ureducerbar bakgrunns horisont som bør tas i betraktning når man diskuterer felleshandling, felles oppmerksomhet og sosial ontologi. Ved å forene Moran og Taylors innsikter om intersubjektivitet, språk og *kommunion* eller emosjonell binding gjennom felles oppmerksomhet, med planleggingsteorien om delt handlekraft fremsatt av Bratman, fremskaffes en adekvat og dekkende redegjørelse for kollektiv handling og samarbeid. I Taylors teori er språket i seg selv det mest grunnleggende nivået som omfatter alle former for kommunikasjon fra det kroppslige til det konseptuelle og intellektuelle. Jeg konkluderer med å hevde at språkets matrise er samtale. Dette forblir slik gjennom menneskets historie og er grunnleggende for all fellesaktivitet og gruppeintensjon.

Introduction

What does it mean to perform an action together? Can our actions be explained by referring solely to our individual intentions, desires, and beliefs? Or do we act together because we are part of a larger whole, a community, that cannot be reduced to the individual members? An important subject within the interdisciplinary field of research on collective intentionality is that of joint action and cooperative activity. How are we able to come together to cooperate, to share intentions, goals, plans and purposes? It seems that the participation in shared practices and social activities is fundamental to understanding how we construct our shared social reality. But what makes these shared practices possible; what characterizes our capacity to engage in a conversation, in an opera performance, or in the joint cooking of a meal? The capacity to coordinate our actions with those of others is a key to understanding our survival and success, both as individuals and as species, but how can joint action be made sense of philosophically and scientifically?

In this thesis we will discuss some central problems of joint action in connection with questions concerning the structure of social phenomena. What can joint action tell us about social groups? And what can groups tell us about joint action? Are group activities most adequately made sense of by a reductivist and individualist account of group function? Or does it turn out that joint action requires reference to irreducible social structures?

I will first discuss some standard theories that are broadly individualist in spirit, and that can be divided into two main camps: non-reductivists, here represented by Margaret Gilbert (1996, 2006, 2014, 2020), John Searle (1990, 1997, 1998, 2010) and Raimo Tuomela (1992, 1993, 1995, 2007), and reductivists, represented by Michael Bratman (1992, 1993, 1999, 2014, 2022). Secondly, I turn to the criticism of these standard theories which are based on what can broadly be termed “enactive theory”, those of Shaun Gallagher (1998) (2006) (2010) (2017), Valentina Fantasia, Hanne de Jaegher and Alessandra Fasulo (2014), Thomas Fuchs and others (2009). These theorists try to move the discussion away from the more or less closed system of “the mental” by stressing the importance of enactment, of embodiment, of gestures and expressivity in the constitution of the joint attention and action. I will then move on to point out some weaknesses inherent in these “enactivist” approaches, arguing in favour of the importance of the mental in understanding collective action and cooperation. Finally, I will respond to both standard and enactivist theories, by introducing a third approach inspired by Dermot Moran,

Charles Taylor, and the classical phenomenological tradition. The phenomenological approach stresses the importance of larger patterns of meaning and habits in understanding the formation of group intentions and actions. These larger and more fundamental background patterns make meaning, communication and thus, the cooperative activities possible. But what are these patterns? And how do these patterns relate to present- and future-directed intentions, planning, and execution of intentional actions? These are questions I aim to address in the final part of this thesis.

Chapter 1: Mapping the discussion

1.1 On the contemporary field of joint action

Numerous philosophical theories of collective intentionality¹ have been developed in the past few decades in order to offer accounts of shared intentions, group agency and the ontological structure of the social by looking at such phenomena as goals, obligations, commitments, shared knowledge, memory, attitudes, and beliefs and connecting these to the epistemic, cognitive, and ontological conditions of the social. Gilbert, Searle, Tuomela and Bratman, among others, have offered accounts of joint agency and its underlying intentional structure, but there is little agreement concerning the actual nature of the phenomenon.

Discussions of collective intentionality goes back to Wilfrid Sellars' first introduction of the concept of "we-intentions", and Raimo Tuomela and Kaarlo Miller also joined the discussion early (1988). However, it was the article of John Searle, "Collective Intentions and Actions" (1990), which set off the debate concerning collective intentionality. With it followed a whole range of questions concerning social ontology, individualism and collectivism, group minds, and shared emotions.

"We-intentions" and "we-attitudes" are connected to normative judgments which express people's shared attitudes. Group-intentions may be described as attitudes that make sense of the actions of individuals acting in concord; they form part of the individual attitudes, but they are not entirely private, for they involve a shared point of view from which the individuals may make sense of their different roles and actions in the shared cooperative activity. Now, do these attitudes, if they really are shared and irreducibly social, imply that the individuals participate in a shared group mind? Are the persons acting together sharing in a kind of social consciousness? Or would such a concept only be a sort of "mystification" that makes social structures mysterious, when they can in fact be clearly explained and reduced to the attitudes and intentions in the minds of the individuals participating in the group?

Generally, one could say that collective intentionality concerns the power of minds to be jointly directed at activities, objects, situations, state of affairs, matter of facts, goals, or

¹ The labels "the philosophy of social phenomena", "collective intentionality theory", and "social ontology" are used roughly to cover the same domain.

ideals. The phenomenon comes in a variety of forms, joint attention and action, shared intention, shared belief, collective commitment, and collective emotion. It is present in our everyday existence, when a group shares certain attitudes together in a demonstration, when two or more people share sufficient common knowledge to be able to start a conversation, when family and friends gather to grieve the loss of a loved one or celebrate during a festival. The field of collective intentionality is relevant for philosophers, psychologists, and social scientists as it's important role in the constitution of the social world is widely agreed upon.

As already stated, this thesis is a discussion on joint action and questions of social ontology related to it. The concept of *joint attention* will also play a vital part, particularly in the last chapter when discussing the work of Michael Tomasello and Charles Taylor. Joint attention is closely related to joint action, insomuch that I will argue that the two activities are inseparable from each other. Briefly, joint attention is here defined as the phenomena of experiencing attention as shared and centred around a common focus. Joint attention gives rise to a fundamental sense of common ground on which agents may be encountered as potential co-operators. Without being able to share attention to the same task, participants would be unable to act together intentionally, which in consequence would bring forth unintentional actions and discordant behaviour.

Joint actions or shared activities or cooperative activities (I will use these concepts as largely referencing to the same phenomena), according to most theorists, must be coordinated and cooperative, achieving certain collective goals, and requiring a coherent formation of shared intentions. Joint attention also presuppose a certain degree of shared belief and common stock of knowledge. Otherwise, the formation of common intentions, on whose coherence the shared focus-point and activity depends, would break down, as underlying background knowledge and shared attitudes are fundamental for there to be a joint activity in any meaningful sense.

Our capacity to participate in joint action is also connected to our ability to engage in joint reasoning and deliberation. Through common action and shared activities and practices we are re-organized and transform ourselves in order that the way we live together, collectively, is in accordance with how we want to live, from our local communities to universalist international politics. For these reasons it is evident that the study of how we share our intentions and attentions in a joint activity, becomes a vital object of study for contemporary philosophy of mind and action.

1.2 The question

Suppose I intend to paint a landscape-painting of the mountain Ulriken, and you intend to paint a landscape-painting of the same mountain. Is the fact that we both intend to do the same thing enough to say that we are painting this landscape together? It seems not. We intuitively know that something else and more is needed for there to be a genuine case of shared cooperative activity. So, what sort of social structures must be in place for there to be a genuine case of joint action? And what is the nature of these “structures”? Firstly, there is joint intention and common knowledge. If I know that you have a future-directed intention to paint Ulriken, I can *say* that “we intend to paint Ulriken in the evening sun,” but the action, when carried out, would still not be joint. The fact that we have common knowledge that “I am going to paint, and you are going to paint” and that we both possess intentions to paint, doesn’t mean that we are going to paint *together*. Perhaps you are going to paint the mountain from the west-side, and I am going to paint the mountain from the east-side. We are doing nothing together even though we know that both are painting and painting at the same time.

So, it seems that aggregates of individual intentions to do something, combined with common knowledge and mutual beliefs, and similar future-directed plans, do not necessarily create collective actions. But what does it mean that you and I are *sharing* in an activity? What is distinctive about “sharing” and “togetherness”? Are collective intentionality something special and irreducible, which makes the activity common and not individual? Are the participants connected to something external to their own individual brains, or is social activity located solely in the individual’s inner workings? This last option, generally known in different versions as the “Individual Ownership Thesis”, claims that each individual has a mind of their own and has a kind of intentional freedom. This claim is thus generally individualist in spirit and stands in opposition to views that understands individual minds as in some way fused when intentions and actions are shared by a group.

To answer these kinds of questions, we have two main camps: reductivist approaches and non-reductivist approaches. To understand what reductive/non-reductive means in this setting, we will look at the following sentences: “We built a house,” “They played the symphony,” and “The team will play tomorrow”. If the plural “We” in “We built a house” signify something more than aggregation of “I”s, then this more might be something non-reductivist, something that cannot be reduced from the group level to the individual level. Is there, in other words, a special “we-intention” operating here that cannot be reduced to the “I-intentions” of the group participants? When several musicians are participants in the

performance of a symphony, is it possible to explain the sentence “they played the symphony” by solely looking at each musician’s individual intentions to play their own notes and do so in coordinated action with the others? If this is so, then one has roughly given a reductionist account. The sentence “the team will play tomorrow” describe a situation that is planned to take place in the future. Do we best make sense of future-directed intentions in a reductionist manner, or introduce a non-reducible notion of the team’s “joint commitment” to act in the future in the manner?

Non-reductionist thinkers like Gilbert, John Searle, and Raimo Tuomela claim that while there indeed are some intentions that are individualistically reducible, there also exists a peculiar set of shared intentions that, at least conceptually and nomologically, cannot be so reduced (2020, 25). For example, if “they played the symphony” cannot be explained solely by looking at concepts and features of the psychological outlook of the individual, but we must introduce some form of irreducible new element to describe the shared structure that enable these musicians to play together in a coordinated manner, then we have posited a non-reductionist analysis. This set of shared intentions is irreducible because it can be conceptualized in terms of certain characteristic properties or states. The properties or states that characterises the orchestra acting as a whole are, according to the non-reductionist account, different than the properties or states of any of the individual musicians and require its own set of concepts for adequate description. In “Non-Reductive Views of Shared Intention” (2020) Tuomela argues that collectively constructed group agents, described as social systems, in many cases seem to be ontologically emergent. That is, they seem to involve qualitatively new features as compared with the individualistic basis, and in this sense these features are irreducible to the individualistic properties of our common-sense framework of agency and persons (2020, 25). Collective ‘states’ in these instances seem not to be reducible to individualistic ‘states.’

For the reductionist accounts, one need not posit such special forms of intention in order to explain the behaviour of a group of people acting together. According to this line of argument, shared intention will be best explained in terms of the properties and concepts already identifiable in our understanding of individual intention and action. One of the most famous and most developed theories in this camp is that of Michael Bratman, for whom individual future-directed intending or “planning” is at the heart of human agency and when combined in the right way does not need any further special concept of an irreducible “we-intention”.

1.3 Action and intention

In Michael Bratman's *Intention, Plans, and Practical Reason* (1999) he defines action, in the context of discussing rational limited beings such as us, as being intentional and directed in some way. This means that actions in some sense must be connected to intentions or the intentional character of human behaviour. I agree with Bratman's definition of intentions as distinctive states of mind and thus irreducible to the desires and beliefs of a rational agent. In defining intention in such a manner, I follow Bratman's deviation from Elisabeth Anscombe and the early Donald Davidson's conceptions of intentions which on the contrary denied that they were distinctive states of mind.² Actions, for Bratman, are done intentionally, or with a certain intention, and actions and intentions are connected in a complex manner. The state of intending to act can be connected to the mind having intentions concerning the future. When I say I am painting a portrait I no longer *intend* to draw it, I *am* drawing a portrait. But intentions need not always be future-directed; there exists both present-directed intentions, intentions concerning what to do beginning now, and future-directed intentions, intentions to do something in a time that has not yet arrived (1999, 4).

For Bratman, we are planning agents which implies that we frequently settle in advance more or less complex plans concerning the future, and then these plans guide our conduct. Our intentions are elements of larger partial plans of action. In consequence, to understand the mental state of intending, it is clarifying to locate it within a conception of human beings characterized as limited intelligent agency, making partial plans and planning important action. My actions can thus be part of a larger plan, a broader more complex form of intentionally formed structure that integrates (ideally) both rational desires and belief in order to anticipate and prepare for future actions (1999, 128).

Let us then return to some real-life scenarios: When two girls sing a duet together, or when surgeons and nurses come together to cooperate in a complex surgery, are there some form of irreducible social phenomena, some "we-intention" underlying the joint action? Are there social structures, beyond individual intentions, desires and beliefs, that enable us to cooperate and to each play our own role in the activity? If two girls sing a duet together, is the sum of their individual actions enough to explain the joint action? Is the sense of "togetherness" reducible to individual intentions, actions, and wills? In response to these and similar

² In his early 1963 paper, "Actions, Reasons, and Causes", Davidson did not see any conceptual need for a concept of intending distinct from those of desire and belief (Bratman 1999, 132). Later, in his 1978 paper "Intending," Davidson changed his mind.

formulated problems, we again meet the two main camps: the reductive and the non-reductive account of group intentions. In the following, we will address these theories.

Chapter 2: The standard theories

2.1 The non-reductive accounts of shared intentions and joint activity

To approach the common accounts of non-reductive shared (or joint) intentions and action, here represented by Searle, Gilbert and Tuomela (although Tuomela have also explored a more reductive thesis³), it is helpful to contrast these against Bratman's reductive "continuity thesis". According to Bratman, joint action is based on the conjecture that the conceptual, metaphysical, and normative structures vital for "modest sociality" (small-scale cases of social activities like cooking a meal or singing a duet together etc.), are continuous with structures of individual planning agency (Bratman 2014, 8). This means that there is no need for a concept of a special irreducible attitude that only applies for collective actions.

Gilbert, Searle, and Bratman agree that there is something significant involved in cases of modest sociality. They can be said to share the belief that small-scale cases may be used as a basis for understanding larger more complex social structures like institutions, political communities, and corporations. In contrast to Bratman, both Gilbert, Searle, and Tuomela see the step from individual to shared agency as involving a new basic practical resource. For John Searle, as we shall see, one need to conceptualize collective action and behaviour by positing a special attitude of the "we-intention". According to Margaret Gilbert, what is needed is a new relation of "joint commitment" between the participants, a relation that necessitates distinctive mutual obligations. Shared intention and activity is for her tightly connected to interpersonal obligations and practices of holding accountable, unique connections binding two or more persons into a unity, as a body and a "plural subject". These connections cannot, according to the non-reductivists, be appropriately accounted for by Bratman and the reductivists which in a "singularist" manner explain such connections solely by concepts that are always applicable to the rationality, intentionality and agency of the individual (Alonso 2020, 42).

What is common to all the non-reductivists, then, is the introduction of a new element on the social level which is not just a matter of common knowledge; it is a new "primitive" irreducible element in our shared agency. Further, they have in common understand similar understanding of larger institutions involving the new element that they consider central to small-scale shared agency.

³ See Raimo Tuomela and Kaarlo Miller's "We-Intentions" (Miller 1988).

When observing the way people think, talk, write, and conduct themselves in communities and in society at large, there indeed seems to be something irreducibly shared and social. And, if one follows Gilbert, we implicitly or explicitly engage in something like a joint commitment when engaging in group activities (2020, 138). A certain kind of expectation and accountability seems to emerge once we, for instance, have consciously commenced singing a song together. Let us say two friends are present at a party and one of them, in a joyous mood, starts singing a melody that she knows her friend knows very well and that they both like. It's a spontaneous action, and yet it is also an invitation to the other of engaging in a group activity. Imagine, further, that the other friend starts singing along. They are now both engaging in an intentional joint activity and also in a form of emotional-bonding by emphasising something they both enjoy and love, as a mood and a style. Even in this spontaneously commenced action in concord, there arises, if we follow Gilbert, an implicit joint commitment and accountability. Now, if the friend suddenly starts singing another song, unknown to the other, she will break with the expectation and obligation to sing with her friend. She will break the joint commitment, the *obligations to one another* to conform to the singing together and bonding together. This obligation is not severe in this case, breaking up the joint activity might not issue any response of blame. The obligations need not, in this sense, imply moral severity; perhaps the response would be one of surprise "why did you suddenly start singing on that one?" or maybe laughter. Still, by breaking the implicit obligations, the other friend is necessarily breaking the intentional joint activity. And so, joint commitments seem to be central to any kind of modest group action.

This is largely how Gilbert thinks about joint action. Following her technical terminology, two or more people constitute a "plural subject" if and only if they are jointly committed with one another in some social setting of shared practice and cooperative activity (2020, 136-37). When persons come together and share in commitments to perform a certain action, one can appropriately say that they are acting as "a body", or a "plural subject," and thus, it becomes proper to refer to them with the collective "we" (2020, 136-37). Notions of non-reducible social phenomena, such as "plural subjects" or "body", are, however, still discussed by the mentioned theorists within a broadly "individualist" framework. They do not endorse the, as they see it, metaphysically suspect ideas of a collectivist "group mind" or "subject" of consciousness "over and above" the individual minds (2020, 137).

Why is modest sociality important? Why start from small-scale cases and build a social theory on this basis? For Gilbert, an advantage of doing so is that it allows for simplicity and

clarity. It becomes easier to “demystify” (a term frequently used by Bratman⁴) the social structures when explaining them using the most central elements that make these structures what they are. By examining simplified situations and constructed schemes, the details of the membership relationships and the action are more easily understood (2006, 97). Another important aspect is that these theories show that matters of *morality* are not always present in cases of mutual obligation and joint commitment. We need to distinguish between *moral* obligations and group obligations as such. To this I will also add that focusing on small-scale cases helps us avoiding “over-interpretation” of social reality, that is, “discovering” connections and structures in social cases that aren’t really there. We will, however, in chapter 4 and 5 see that the modest sociality-approach might nevertheless be questioned from the standpoint of a more holistic phenomenological standpoint.

We shall now return to the contrast with Bratman, who will not use concepts such as “joint commitment” in a non-reductive sense, even though he endorses a singularist or continuity-based version of commitment and obligations in the interrelations of cooperative activity. While Bratman invokes concordant or “meshing” inter-personal intentions and thus concordant personal commitments, he differs from Gilbert in that he never makes commitment itself the most central basis for sociality.

Bratman explains group activities by looking at the intentions and attitudes of the individual participants in the group partaking in the action. Why should a non-reductivist account of group activity be more accurate? Arguing for discontinuance between the individual and the group level seems plausible, because it makes better sense of the unique experience of “unity” often associated with social groups. The appropriateness of using the collective “we” in “we are playing jazz” or “we are dancing together” seems to require that we introduce new concepts that are only appropriate at the social level. Also, trying to reduce shared intentions and actions to individual intentions and actions, seem to introduce an unnecessary “roundabout” way to explain joint action in cases where it can be adequately and sufficiently explained on the social level without needing to “weave” individual intentions into a complex web in order to account for the same.

To further test the effectiveness and appropriateness of the non-reducible accounts, such as Gilbert’s, I will introduce another case. Suppose that Peter and Hallvard have jointly committed to play an online computer game. In this game they each play an “avatar” participating in the activity of defending an area from enemy online players. Peter ought now

⁴ See for example *Intention, Plans, and Practical Reason* (1999, 167) and *Shared Agency* (2014, 25)

to play a defender of the area. Suppose now that Peter suddenly decides to leave his post of defence, fleeing the area that is supposed to be defended by his online avatar. It seems that Hallvard (and the other online members of the group) is in a position to question Peter's action on the basis of the group's joint commitment to act as a plural subject, to act as defenders from enemy attacks. Furthermore, Hallvard is in a position to call Peter back to his post. And should the case appear to be the other way around, should Hallvard intentionally abandon his posts, then Peter and the other members will be in a position to call Hallvard back to his post. Here the joint commitment seems to be fundamental and collectively shared, not by any one person, but by the "plural subject" the "we" that binds the different individuals together in a shared activity. The parties' joint commitment to accept as a body the relevant goal, would explain these observed phenomena. It is by no means clear that anything else, a reductive account say, can give an equally satisfactory explanation.

In a similar manner to Gilbert, Searle argued that the intentional behaviour of groups "is a primitive phenomenon that cannot be analysed as just the summation of individual intentional behaviour" (Searle 1990, 401). The difference lies in the understanding of the nature of the irreducible social phenomena. For Searle it is the capacity to form the "we-intention", rather than joint commitments or plural subjects, that forms the fundamental basis for the social. Collective intentions expressed in the form "we intend to do such-and-such" or "we are doing such-and-such" constitute a primitive phenomenon that cannot be analysed in terms of individual intentions expressed in the form "I intend to do such-and-such" or "I am doing such-and-such" (1990, 401).

Collective intentions are understood to be a biologically primitive sense of the other person as a candidate for shared intentionality. It is what makes sociality possible, what made us able to construct the social world. Thus, it is an ambitious concept, meant to connect our contemporary physics and metaphysics of the natural science with a social science that can make biological sense of the creation and maintenance of social norms, conventions, and institutional facts, such as elections, marriages, business corporations, property rights, nation states etc.

In contrast to social coordination, which is not completely unique to humans and can be found among flocks of birds, tribes of lions etc., the capacity for "we-intention" is a unique form of shared or collective intentionality that enables highly conscious joint action. Hammering out a statue or constructing a public marketplace in a city requires more than social coordination; one need a strong sense of the "we" and a strong sense of a common goal or good. Searle's view is still "individualist" in the sense that we-intentions are attitudes in the head of

an *individual*, even though they concern the activity of a supposed “we”. You could have a we-intention, in Searle’s sense, if you were the only person in the world, but thought there were others with whom you could cooperate with. Searle himself is deeply sceptical of philosophers that embrace ideas such as group minds or collective unconscious; he even ridicules them (1990, 404).

To what degree may we say that Searle’s we-intentions are similar to joint commitment? The we-intention is quite easy to activate by our biological capacity; when engaging in some group action, we will experience the sense of the “we”. In this sense it is close to Gilbert’s plural subject held together by joint commitment which is also activated as soon as several individuals coherently participate in a group activity. The human species seems to possess a capacity for other agents coming together and forming a “we” and an “us”, and this is underlined in both Searle and Gilbert’s analysis.

Cooperation requires more than just an aggregate of individual “I-intentions”, according to both Searle and Gilbert. It is possible for many individuals to have a goal, to know that the other individuals have the same goal, and to share with the others a common knowledge that everyone have the same goal, without there being any cooperation taking place. An irreducible and socially shared “directedness” of the mind, a collective intention, seems to be necessary for a genuine cooperative activity to take place. Four people might all turn up at the same café around lunchtime, they might all know that the others are planning on eating lunch, and yet they might not at all share a meal *together*. A shared sense of the “we” might be entirely absent; they might even dislike each other and eat at tables as far away from one another as possible.

Let us turn to Tuomela’s non-reductive account. Which are-, and how do we understand, the unique properties of social groups and cooperative activities? Following Tuomela, the individual and the social levels exhibit two different sets of properties. On one hand we have individualistic properties, the “I-mode” or “private” mental states of the group members. On the other, we have a set of group-level properties that becomes apparent when groups do something together. Group-level states or properties could be said to *supervene* the set of individualistic private properties. This supervenience is possible if and only if for each change in the collective state there is some change in the individualistic ones (Tuomela 2020, 26).

The group level intentions are irreducibly shared, but are dependent on the intentions of each group member. A group property cannot change on its own. It has no independent causal nexus and must always correspond to changes taking place in the participants’ relevant individualistic single-agent or multi-agent properties (2020, 26). The we-mode, broadly the perspective one takes as a member of a group, is contrasted with the “I-mode,” the perspective

one takes as an individual. This contrast is designed to differentiate the personal, or “private”, from the social, and hence to provide a distinguishing characteristic of the social. Although reasoning, thinking, intending, or believing from the we-mode is something individuals do, the “we-mode” cannot be reduced to the I-mode.

Typical examples of groups that seem to display such supervenient (but still irreducible) properties are business corporations and public institutions: “We” are corporation that aim at excelling in the production of “such-and-such”, “we” contribute to deliver “such-and such” solutions to society, and “our” core values are “this-and-this”. In the case of a corporation, no one of the individual members need to share the public values and goals of the corporation. One can imagine cases where all individuals privately deeply dislike the official values and images of the corporation they are part of. Thus, corporations and institutions can be viewed as hierarchical structured groups and their parts with the shareholders/administrative office-workers/bureaucrats at the top level. This is what non-reducibility amounts to in Tuomela’s discussion. If one does not accept a reduction of these social structures, then one ends up one form or another of these “we-intentions” or shared irreducible supervenient intentions (2020, 26).

Tuomela’s point about larger social structures is similar to Gilbert’s theory of the “plural subject” which is centred on the idea that shared intentions are irreducible at least in part because a shared intention does not really require corresponding personal intentions by the participants (2020, 27). The “we” in a shared intention case is non-distributive rather than distributive. What does this mean? Gilbert lists three criteria for an adequate account of shared intention: the disjunction criterion, the concurrence criterion, and the obligation criterion. The first, *the disjunction criterion*, states that an adequate account of shared intention must be of such a nature that it is not necessarily the case that for every shared intention there exists a correlative personal intention of the individual parties (2020, 21). This means that in joint action, the content of a personal intention need not be the same as that of the shared intention of the group; the collective intentions cannot be reducible to personal intentions due to the simple fact that the latter might not even exist. What “we” as a plural subject intend by our joint action need not correspond to what the individual “I” intend.

Gilbert’s second criterion, *the concurrence criterion*, maintains that for an account of collective action to be adequate, it is necessary that, absent special background understandings, the concurrence of all parties is required in order that a given shared intention be changed or rescinded, or that a given part be released from participating in it. There must, in other words, be a shared consensus among the participants for there to be genuine shared intention to act or

a change of intention to act. The third, *the obligation criterion*, states that a shared intention implies that each party to a shared intention is obliged to act as appropriate to the shared intention in conjunction with the rest. That is, there needs to be obligations centred on the joint action that keep the group together and implicitly hold the group members accountable and reliable. This does not imply that such obligations need necessarily be of a moral nature: Persons have reciprocal obligation to one another to perform certain actions, and thus they owe one another these actions, making everyone accountable if they do not act out their role (2020, 27). On these three criteria Gilbert formulates her *plural subject* account of shared intention: 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 (2020, 27).

For Margaret Gilbert, then, the shared intentions of a plural subject essentially involve a joint commitment. We intend to J if and only if we jointly commit to J-ing as a body. Joint commitments are formed when each individual expresses his or her willingness to be so committed with others. Joint commitments are not individual commitments to do one's part. Rather, a joint commitment is the commitment of a group. It cannot be dissolved without the agreement of all parties and the existence of joint commitments bring about obligations and entitlements.

If we turn to Tuomela we find a different analysis of joint intentions which emphasises the togetherness of each individual within a group in possessing the right kind of "we-intention":

A member A_i of a collective g we-intends to do X if and only if: (i) A_i intends to do his part of X (as his part of X); (ii) A_i has a belief to the effect that the joint action opportunities for an intentional performance of X will obtain... (iii) A_i believes that there is (or will be) a mutual belief among the participating members of g ... to the effect that the joint action opportunities for an intentional performance of X will obtain... (iv) (i) in part because of (ii) and (iii) (2005, 340-41).

This amounts to stating that individuals must have the right "mutual beliefs" about the beliefs of the other members of the group for there to be, in a strict sense, shared intentions directed towards action (Chant 2020, 17). Individuals have the shared belief that such and such is the case if everyone in the group believes such and such is the case, and so on. In everyday life, this would be the "common knowledge", a basic shared understanding of what everyone intends

and that everyone intends it in common. Furthermore, the participants need to be in possession of a we-intention to do their part *as their part* of the collective action (2020, 17).

In the above, I have described the non-reductive accounts of shared intention and action. There has to be, according to these approaches, an irreducible we-intention, or irreducible plural subject constituted by joint commitment, or *supervening* “we-mode” made possible by common knowledge and the right kind of overlapping intentions, in order to genuinely achieve joint action. Individuals accordingly participate in a joint action because they all have intentions of the form “we intend to J”, and the capacity to form such intentions is something unique, something that cannot be reduced to an aggregate of individual intentions. This capacity presupposes other background capacities, other conditions necessary for certain cognitive activities and language. It also presupposes a background sense of the other as a candidate for cooperative agency.

The non-reductive theories of Gilbert, Searle and Tuomela, may however exhibit some weak points. We might agree with the non-reductivists about the sense of something special and irreducibly being present in the social, but does this necessitate a biologically primitive capacity for “we-intention” or a “plural subject” in every case of modest sociality? And how exactly are all the different individuals in a group able to interweave their respective “we-intention” or “joint commitment” or “we-mode”? As every individual is different, there is little reason to believe the we-intentions are the exact same ones? If there are as many “we-intentions” as there are individuals, then how can they be at the same time irreducibly collective and shared? If they have other properties and states than individual properties and states, how can they nevertheless be coordinated and harmonized by the individuals? And if the shared “we-intention” constitute one “whole”, then how can we avoid breaking with the “Individual Ownership principle”, the principle that there are no states or intentions over and beyond individual minds guiding joint action, which seems to be fundamental for at least most of the standard theories? If there is one shared intention that structure the activity “over and above” the individuals that participate in it, then it seems at least unclear how this supervenient structure at the same time is connected to a biological primitive capacity in an individualist way. In the following, we will address the reductivist theorists to see if this approach might make better sense of joint action and social ontology.

2.2 The reductive theory

Following the reductive line of arguments, shared intention are best explained in terms of concepts and features already established in our understanding of individual intention and action. One of the famous and most developed theories within this camp is that of Michael Bratman, for whom individual (future-directed) intending or “planning” is at the heart of human agency and a special concept of an irreducible “we-intention” is redundant.

Crucial to Bratman’s theory of modest sociality (small-scale cases of social activities and group action), is the idea that *shared intentions* interpersonally structure and coordinate thought and action, and this involves associated social norms (2014, 33-34). Our intentions have different roles, and some are long-term and future-directed intentions which put constrain on our daily habits, goals, options and on present-directed intentions. Let us say that my brother and me are going jogging tomorrow. In this case, if we are rational, we need to make sure that our future-directed intentions intertwine with one another in concord. I must contact my brother and form a plan; we need to settle on a long-term future-directed intention on what we are going to do and when. Our deliberations will result in a plan which must follow the appropriate set of norms that make sure we can coherently act on it. Within this larger plan, “smaller” means-directed intentions are embedded: perhaps I intend to get a laundry done today so that I have clean training-clothes for tomorrow. Perhaps my brother intend to talk to his wife today in order to make sure that his jogging-plan tomorrow will not conflict with her plans. My intention to do the laundry and his intention to talk to his wife are both “means-intentions” embedded in the larger plan to an end, which is to go jogging tomorrow. So, our intentions must be guided by norms of *means-end coherence*. And if we are to execute our plan, and commence the joint activity of jogging together, we must also make sure, not only that our general plans to jog intertwine, but also that our “sub-plans” mesh such that no conflict or misunderstanding ensues. Furthermore, if my brother only plan to go jogging in order to humiliate me, by showing off his superior perseverance and then run away from me, the activity would not be joint activity in the planned sense. So, our intentions need to be guided by norms of *stability and consistency*. Notice that in this example, what makes the cooperative activity (jogging) possible, is the intertwining of the individuals’ plans and subplans, guided by rational norms of planning, *not* our participation in some special collective we-intention. At this point, we haven’t needed to introduce any social structure beyond the coordination of individual plans to act together.

Unlike the non-reductivists, then, a planning theory like Bratman’s plays a central role in explaining interpersonal actions and activities (2014, 8). Planning is key to some central

characteristic roles because it enable us, limited rational beings as we are, to organize and coordinate our actions in time. Intending to perform an action in the future means to take the action as an orientation point, either as a means or as an end, in our further deliberations and to reason in certain ways. Thus, the goal of the reductive theory is to make a sufficient description of the group activity without reference to a higher “we-mode” of group reasoning. It aims to show that if the plans and subplans of one member of a group can be coordinated in the right way with the plans and subplans of the other members, a further reference to an irreducible social attitude or structure is not needed. As long as the individual is “doing his or her part” in the activity and the plans and subplans harmonize, it seems to be possible for a group to realize a collective goal.

Now, this ambition of the reductivists can be challenged on the grounds that the notion of an individual’s “doing his part toward achieving the *collective* goal” seems to introduce an element of circularity in the description. If I am doing my part in the activity of jogging such that I can say that “we are jogging”, have I not implicitly included a notion of we-intention or at least some form of social attitude that “doing his part” fits into? As long as the individual intends to partake in the larger pattern of collective activity, the concept “collective” is either explicitly or implicitly presupposed. And if we interpret a notion of “doing his part” as not referring to a collective goal of performing a certain activity together, then the analysis will be too weak; what would it mean to plan to do something if the plan and intentions were not in some sense shared?

Bratman tries to solve this problem by arguing that shared intentions underlying joint action can be understood as a state of affairs consisting primarily of individual intentions of the form “I intend that we J” (where J is a joint action) and mutual beliefs regarding these intentions (1993, 99). In order for there to be something specific about collective intention and activity without reference to a particular “we-mode,” the notion of meshing *subplans* becomes vital. What distinguishes shared activity is not only a proper kind of shared planning, that is individual future directed intention to perform joint activity, the members of the group must also have different private motivations and subplans, that don’t create discord with the subplans of the others. Bratman’s analysis can be restated thus:

We intend to J if and only if:

- (1) (a) I intend that we J; and (b) you intend that we J.
- (2) I intend that we J in accordance with and because of 1a and 1b, and meshing subplans of 1a and 1b; you intend the same.

(3) 1 and 2 are common knowledge (1993, 99).

The requirement that subplans mesh is meant to rule out instances where participants both intend to do some joint action, but they intend to go about fulfilling that intention in ways that would ultimately undermine the activity itself. Meshing subplans, in other words, requires that the means I use to complete my part of the joint action will not prevent you from completing your part of the joint action.

Thus, for Bratman, shared intention can be reduced to a complex structure of attitudes of individuals. It involves intentions of each individual in favour of the joint activity itself—where the latter is understood as neutral with respect to shared intentionality. When you and I share an intention to jog together, this means that I intend that we jog, and you intend that we jog. Unlike Searle’s we-intention, Bratman argues that joint activity is not a special type of intention in this case, but simply a conflict-free intertwining of individual intentions, that is, the intention to jog together. Does this mean, then, that there is no intention to do something in common? Are these intentions examples of intending in favour of the joint activity only in name? According to Bratman, what an individual intends in shared intention is, strictly speaking, the joint activity itself (Alonso 2020, 38). In addition to this, the intention necessarily involves some cognitive interdependence between the attitudes of individuals usually in the form of mutual knowledge, beliefs, or assumptions about each other’s intentions, although an extra we-intention is not needed. Cognitive interdependence, for Bratman, makes it possible for individuals to coordinate their thought and action in ways that can be described as shared intention, for it allows each of them to plan and act on the assumption that the others have the right attitude and will perform the right actions when the time comes.

Such intentions are, according to Bratman, necessarily “reflexive” and “interlocking”, that is, each member intends that both his own intention and the intention of the other be effective. They track the concord (what Bratman calls “meshing”) of lower-level intentions (what he calls “subplans”) of each, concerning ways of carrying out the joint activity. Everything, then, depends on whether individual intentions can be coordinated in such a way that interdependent activity becomes possible (2020, 38). In the article *Shared Cooperative Activity* Bratman explores the fundamental structures that make possible coordinated cooperation, which is a more concentrated and specific form of activity than simply having shared intentions in a specified sense. The concept of a shared cooperative activity points out a distinctive type of interpersonal interaction which is fundamental to joint acts such as the playing of music, partaking in a conversation, or rowing together.

Shared cooperative activity (SCA) involves, according to Bratman, appropriate behaviours; If you and I successfully engage in the cooperative activity of rowing together, then it's crucial that we perform the proper tasks appropriate for the activity of rowing. But we might row together without acting cooperatively. Perhaps neither of us even knows of the other's activities, or though we each know of the other's activities, neither of us cares. Given these sorts of situations, what is needed for there to be SCA? Bratman supplements the argument with a hypothetical situation:

Suppose that you and I sing a duet together, and that this is a SCA. I will be trying to be responsive to your intentions and actions, knowing that you will be trying to be responsive to my intentions and actions. This mutual responsiveness will be in the pursuit of a goal we each have, namely, our singing the duet. You may have this for different reasons than I do; but at the least we will each have this as a goal. Finally, I will not merely stand back and allow you to sing your part of the duet. If I believe that you need my help I will provide it if I can (1992, 327-28).

Bratman uses this hypothetical duet in order to help us identify a trio of features characteristic of shared cooperative activity: (i) *Mutual responsiveness*: In SCA each participating agent attempts to be responsive to the intentions and actions of the other, knowing that the other is attempting to be similarly responsive. Each seeks to guide his behaviour with an eye to the behaviour of the other, knowing that the other seeks to do likewise. (ii) *Commitment to the joint activity*: In SCA the participants each have an appropriate commitment (though perhaps for different reasons) to the joint activity, and their mutual responsiveness is in the pursuit of this commitment. (iii) *Commitment to mutual support*: In SCA each agent is committed to supporting the effort of the other to play her role in the joint activity. If I believe that you need my help to follow the joint rhythm for the rowing exercise, I am prepared to provide such help; and you are similarly prepared to support me in my role. These commitments to support each other put us in a position to perform the joint activity successfully even if we each need help in certain ways (1992, 328). It is evident that there may be social cases where all of these features are not present; in cases of enmity, war and rivalry, mutual responsiveness might not imply commitment to joint activity and commitment to mutual support. In the familiar spiral of reciprocal expectations, enemies will do their best to be responsive to the other, and each will take it for granted that they are playing or acting out in a field of action, but they will not cooperate. On the contrary, cooperation would have made them in some sense allies. So, there

can be mutual responsiveness in the pursuit of personal goals without commitment to a joint activity and without commitment to mutual support (1992, 329).

Bratman further enquires into the nature of *commitment* in joint activity. What is it for two persons to commit to the joint activity of a “we”? His initial conjecture is that this commitment typically involves, in part, an *intention in favour of the joint activity*. The reasons for joining in may be different, but there needs to be a motivated reason and consciousness in order for there to be participation at all. Here Bratman distinguishes what he calls joint-act-types that are *cooperatively neutral* from those that are *cooperatively loaded* (1992, 330). A cooperatively loaded joint-act-type—for example, trying to solve a problem together—already brings in the idea of cooperation. In contrast, in the case of cooperatively neutral joint-act-types, joint performance of an act of that type may be cooperative but need not be. We might perform many shared tasks together, going out for a promenade, painting a house together, which need not involve cooperation. The behavioural conditions, as Bratman calls them, for SCA, are more specific accounts of what intending together means. According to Bratman, SCA involves, crucially, the roles of future-directed intentions as elements of partial plans. This planning conception of intention allows us to be more liberal about what can be intended than about what can be attempted; for references to things other than our own actions can function appropriately in our plans. I engage in planning aimed at settling on means to our joint action: I can, for example, figure out how to support our singing the duet—perhaps by helping you find your music sheet. And I can try to ensure that the rest of my plans are consistent with our performance of a certain joint action: I can, for example, eschew ways of singing that will prevent your coming in on time. In these ways the planning conception of intention supports the legitimacy of the appeal to my intention that *we* J.

In our SCA, then, you and I each intend that *we* perform the cooperatively neutral joint action (1992, 331). But this is not enough or completely appropriate to ensure the correct kind of commitment to the joint action. What about cases of meshing subplans and interdependent intentions where subplans conflict with each other? Let us say that our individual subplans concerning our J-ing *mesh* just in case there is some way we could J that would not violate either of our subplans but would, rather, involve the successful execution of those subplans (1992, 332). When our subplans mesh, we can have different plans, but successfully carry out the cooperative task, whereas when our subplans do not mesh, there is a conflict making SCA impossible.

This suggests that in SCA each agent does not just intend that the group perform the cooperatively neutral joint action but also intends that the group perform this joint action in

accordance with subplans (of the intentions in favour of the joint action) that mesh. Concord between the different subplans of the participants is necessary for the carrying out of intentional joint action. We must therefore, according to Bratman, build a meshing condition into the *content* of each individual's intention. If members in a group performing a shared cooperative activity do not consciously maintain the mesh, the concord, of private planning-intentions (subplans), we risk not being able to collectively carry out the action if one or several members should decide to change the content of their subplans. If several members change the content of their subplans, the concordance might be lost, and the cooperative attitude characteristic of SCA will be absent (1992, 332).

Some intentions that do mesh might still not be accepted, of course, if they are beyond the pale. What the group will be prepared to accept will depend on our relevant desires and intentions as they relate to particular practice or activity. Thus, according to Bratman, in shared cooperative activity each agent intends that the group perform the joint action in accordance with subplans that mesh (1992, 333).

Collaboration is cooperation between intentional agents each of whom sees and treats the other as such. For our being able to perform our joint action, for this to be SCA, the one must intend that we J (do the joint action) in part *because of* your intention that we J and its subplans. In this way my intention is depended on- and favours your participation as an intentional agent. In SCA I will see each of the cooperators as participating, intentional agents (1992, 333).

In SCA each agent intends that the group perform the joint action in accordance with and because of meshing subplans of each participating agent's intention that the group so act. That is, for cooperatively neutral J, our J-ing is a SCA only if I intend that we J, I intend that we J in accordance with and because of the meshing subplans of me intending so and you intending so, you intend that we J in accordance with and because of meshing subplans of both of us intending that we J. Our intentions and subplans *need not be fully formed*, according to Bratman, for us to already perform SCA, but if the agents' bargaining does lead to meshing subplans our resulting J-ing may still be a SCA.

It is not that one needs to know very explicitly about these intentions. The agents may neither know nor care about the details of the other's subplans, as long as they, in reality, mesh. According to Bratman SCA cannot include cases where the intentions of the participants are coerced by the other participants (1992, 335). Shared cooperative activity must be voluntarily performed. In SCA the fact that there is this mutually uncoerced system of intentions will be a matter of common knowledge among the participants. I will know that we have these intentions,

you will know that we have these intentions, I will at least be in a position to know that you know this, and so on. It is the web of intentions that ensures the commitment to the joint activity characteristic of shared cooperative activity. The system of intentions characteristic of SCA must be *interlocking*; for each agent must have intentions in favour of the efficacy of the intentions of the others. The intentions must be mutually end-providing, reflexive, and self-conscious. SCA involves appropriately interlocking and reflexive systems of mutually uncoerced intentions concerning the joint activity (1992, 336).

In addition to this, we remember, the third feature of shared cooperative activity is the commitment of each agent to support the other's attempts to play her role in the joint action. So, I am committed to helping you play your role in our joint action to the extent that I believe such help to be necessary. Let us consider another example: You and I are engaging in a couple dance. But I have no disposition at all to help you follow my steps, should you stumble or fall out of the rhythm; for I prefer your failure to our success. Were you to unexpectedly stumble, I would gleefully allow you to be embarrassed in front of the audience—as I might say “One wrong step and I’ll abandon you to the wolves.” And you have a similar attitude: you fully expect me to get my movements correct, and so you intend to move in a coordinated way that meshes with mine. But were I to stumble, you would not help, for you prefer my failure to our success. We each intend that we perform the couple dance as we expect it to be, and we each intend that we do so by way of meshing subplans. But we do not have *commitments to support each other* of the sort that is characteristic of SCA. If we, as unhelpful dancers, perform the dance together our dancing may be *jointly intentional*, but it is not a shared cooperative activity. Bratman gives the following account of the attitudes involved in SCA:

- (1)(a)(i) I intend that we J.
- (1)(a)(ii) I intend that we J in accordance with and because of meshing subplans of (1)(a)(i) and (1)(b)(i).
- (1)(b)(i) You intend that we J.
- (1)(b)(ii) You intend that we J in accordance with and because of meshing subplans of (1)(a)(i) and (1)(b)(i).
- (1)(c) The intentions in (1)(a) and in (1)(b) are not coerced by the other participant.
- (1)(c) The intentions in (1)(a) and (1)(b) are minimally cooperatively stable.
- (2) It is common knowledge between us that (1) (1992, 338).

In addition to this Bratman also adds the condition of mutual responsiveness of intention, and stresses that a joint activity can be cooperative down to a certain level and yet competitive beyond that. If it is competitive, like in chess or in combative play of fighting sports, there is joint activity but not shared cooperative activity (1992, 340). Joint action and joint intention, in his view, embrace a whole different range of activities compared to the stricter concept of SCA. SCA involves mutual responsiveness—of intention and in action—in the service of appropriately stable, interlocking, reflexive, and mutually noncoerced intentions in favour of the joint activity.

I generally find this way of accounting for joint action and shared cooperative activity more convincing than the non-reductivists' accounts. This is because Bratman tries to understand what distinguishes group activity and shared cooperative activity in terms of the attitudes and actions of the individuals involved. In his theory, he illuminates the depths and complexity of how we are able to act together, and yet shows this without mystifying or making obscure our relations to each other or to the activity itself. Yet, both standard theories, the reductive and non-reductive alike, seem to be missing something crucial. By restricting, in an individualist spirit, their analyses mostly to constructed and abstract cases of small-scale sociality and shared cooperative cases, the standard theories, reductive and non-reductive alike, focus too little on the larger background patterns and embodied meanings of our communal actions. I will in the following move on to some theories that criticize the standard views.

Chapter 3: The enactive approach

3.1 Critical perspectives on the standard theories

Now that we have looked at the reductive theory of Bratman, we have learned about how the complex structure of interconnected individual future-directed intentions make joint activity possible. This analysis of meshing plans and subplans helps to understand the underlying structure of cooperation and participation. But could not this explanation, to some degree, be seen as a way of explaining *away* the actual phenomenon of the social? Is not human rationality always incarnated in human history, at a specific place and at a specific time? Placing all conceptual centrality of explanation in the minds of different individuals, into the individual's interior mental "space", can lead to an underestimation of the importance of embodied interaction and expressive communication when participating in specific communities of shared meaning and shared social imaginaries⁵.

Shared intention, or a common goal, seems to be not just a matter of technical mental schemes in the minds of individuals with access to the plans of other people, but also seems to involve embodying and expressing visions or images of the future. An example could be the cooking of a meal as part of the preparation for a festive occasion where friends and family come together. The participants involved in preparing the food do not only cook because they have planned to do it, in the sense of having found a date and clocked it. They come together because they have a shared future image of the festive moods and rituals that they participate in, through space and time, involving larger patterns of reciprocity, gift-exchange, and narrative unfolding. Intentions are connected to this, but for there to be a shared intention or common goal, such events are driven by both a desire to participate in festive activity with friends, as well as by a background of social imaginary about what is worthy of communal celebration, and song. Bratman's theory of intention, while ingenious in explaining in detail what might be

⁵ The concept "social imaginary" is borrowed from the philosopher Charles Taylor. He writes: "What I'm trying to get at with this term is something much broader and deeper than the intellectual schemes people may entertain when they think about social reality in a disengaged mode. I am thinking rather of the ways in which they imagine their social existence, how they fit together with others, how things go on between them and their fellows, the expectations which are normally met, and the deeper normative notions and images which underlie these expectations" (2007, 171). Taylor's exploration of the social imaginary of modern society can be found in *Sources of the Self* (1989), *Modern Social Imaginaries* (2004) and *A Secular Age* (2007).

going on half-consciously in the minds of a group engaging in a joint activity, gives little weight to the social imaginary dimension and the embodied interactive dimension of everyday participatory activity.

We have seen that in both the reductive and the non-reductive theories, the main trajectory seems to originate from cases of small-scale sociality, or at least they seem to presume a somewhat “disembodied” or “excarnated” philosophical vocabulary⁶ based on constructed cases of group behaviour, in trying to account for joint action and social ontology. Granted, Bratman often emphasises that his planning theory isn’t meant to take into account the entire complexity of the social realm, but only highlights an important aspect of it (2014, 24). However, the conceptual vocabulary of his and the other standard theories nevertheless reflects the somewhat limited social imaginary of a specific philosophical culture. Why, for example, “modest” cases of sociality are assumed to be “gold standard” for explaining any cases, is rarely discussed. Could not a more incarnated, historical, and dialectical approach, addressing the whole range and meaning of collective action, from the simple to the exceptionally heroic and self-transcending, give a more truthful account of joint action? Generally, a vocabulary more informed by the embodied and expressive meanings in our interactive communication often seems to be lacking. In the following, I shall discuss thinkers who defend theories which take the embodied dimension of sociality more into account.

These theories, the enactive theories of Thomas Fuchs and Hanne De Jaegher (2009), Shaun Gallagher (2010), Valentina Fantasia, Hanne De Jaegher, and Alessandra Fasulo (2014), defend the proposition that joint action is based primarily on embodied, enactive, and extended interactive processes. The coordination involved in joint activity and joint attention is primarily constituted by embodied movement, gestures, facial expressions, and communicative practices,

⁶ I’ve borrowed this concept from Charles Taylor. For him, the movement of “excarnation” in modern civilization is the movement away from the lived body, the en fleshed and ritually rich forms of life in philosophy, religion, and culture, into evermore “mind-centred”, anti-metaphoric and conceptually abstract ways of thinking and living. Philosophy and reason becomes disentangled from analogical, narratological and metaphorical ways of perceiving the world. Taylor calls this excarnation the «exaltation of disengaged reason as the royal road to knowledge, even in human affairs. The proper road to knowledge is by objectification, even in history [...] By “objectification”, I mean grasping the matter studied as something quite independent of us, where we don’t need to understand it all through our involvement with it, or the meanings it has in our lives” (2007, 746).

placed in specific historical contexts. In such interactive situations, we co-constitute the meaning of the other person's actions and the world around us (2010, 112).

Let us look at the description set forth in the standard theories, where partners make individual plans for achieving a common goal, and then articulate predictions upon the other's intention to achieve the same goal. Is this description really a paradigm for all group activities? Several of our actions seem not only to be technical problem-solving or aiming at some goal, but they may express certain ways of life, social footings, and shared meanings. Our embodiment communicates through styles and moods of actions, through enactment.

According to the enactive approach, the problem with the standard theories is that they take for granted a "mentalist" understanding of meaning and intention. The standard theories, as they see them, presuppose a "private" understanding of intentions, based on the assumption that we start with a conscious state inside ourselves and then we need to align our inner mental state to the outside, trying to reach out to others with minds like our own. In other words, we start from our inner citadels and *then* ask how we can connect and cooperate with others; we start with the inner "mental" and then reach out to the outside hoping to communicate and form shared intentions and subsequently commence activities together. However, how is it possible to connect our *different* intentions, beliefs and knowledge and form *shared* intentions, beliefs and knowledge?

According to enactivist theory the above construal, is, however, misguided and we shall move on to study Thomas Fuchs and Hanne de Jaegher in "Enactive intersubjectivity" (2009). Their approach is to argue against standard theories in social cognition. Within social cognition there are two dominant theories: *the theory of mind theory* and the so-called *simulation theory*. *The theory of mind theory* (TT) states that we can explain and predict another person's intentions by relying on an innate or acquired "theory of how people generally behave in the world, and how the mental states of human beings, such as beliefs, desires, attitudes cause the behaviour of human beings (2009, 467). Thus, by relying on our theory of how others generally work, we can make inferences about the others' mental states and doing so allows us to act and think together.

Moving on to the *Simulation theory* (ST), the understanding of social cognition is based on our ability to make *models* of things in the world. This theory states that we possess an inner model, in fact our own mind, that we use for simulating another person's mental state (2009, 467). The ST, in other words, is based on the proposition that the beliefs and intentions of others are understandable because we are able to think *as if* we were in their situation, or *as if* we were them (2009, 467). We can simulate their beliefs, attitudes, their intentions in general, and this

allows us to come together to form collective intentions and actions. While it is true that the standard theories mentioned in the earlier chapters base their understanding explicitly on these two theories, the claim of the enactivists is that the underlying picture of “the mental” seems to be taken for granted by most of the standard approaches to joint action and social ontology. This underlying picture is the assumption that thoughts, intentions, and meanings are trapped in the brains of individuals, and that we always therefore move from our mental states in order to reach towards the outside world or the other subjects. Fuchs and Jaegher never claim to give an exhaustive account of this “mentalist” framework in all the standard approaches, but focus on giving a general characterisation of the mentalist paradigm.

According to enactive theory, the mentalist approaches reflect a general problem with their account of cooperation: cooperation is framed in its full-blown, adult form and therefore is not applicable to those who do not possess high socio-cognitive skills (2014, 2). Furthermore, both TT and ST (and similar theories) conceive the mental as a realm separated from others by an epistemic gulf that can only be crossed by inference or projection or some other capacity that the brain possesses. But according to the enactivists, we do not need to look inside ourselves or even be very conscious, to be able to recognise or connect with the thoughts and feelings of others. We only need to observe their faces or their body language (2009, 468). In the realm of everyday life and common practice, we are seldomly in need of any special introspective simulation or inference when we interact with people. When a group is centred around some shared goal or activity, there is usually no need for a theory or a simulation, or even a consciously formed intention. Perhaps aside from a case of some great difficulty or a tough confrontation against someone not trusted, or a rival, or someone who have reasons to hide their intentions. In most cases, we immediately perceive the other’s intentions and emotions in his embodied behaviour as related to a meaningful context (2009, 468).

The other problem of the standard theories is that they tend to treat human beings as if they primarily observe one another from a third-person stance; the paradigm of the inner and private “mental states”. They are in general biased towards localising social cognition in one participant or in his brain. But this, according to the enactivists, does not correspond to empirical reality. Our primary and everyday encounters with people we know, with family, community, and friends, are not solitary observations but interactions in the second-person perspective. Only in special cases, in situations of doubt, suspicion, rivalry and confusion, one might try to make explicit calculated predictions or inferences about the other’s intentions. When we find the behaviour of others in the group ambiguous, when the context is not enough

to immediately understand the situation, then we may find the need to reason carefully about it (2009, 468).

The enactivists claim that traditional theories, both those of social cognitive science and those of collective intentionality, largely assume a disembodied sender-receiver relation between two Cartesian minds (2009, 468). In this schema, the body more or less solely functions as a transmission device, a kind of communicative “avatar”. It is true that some simulation theories increasingly have begun to include the body in the modelling of other minds’ intentions, but even so, they still do not consider the *reciprocity* of embodied agents (2009, 468). So far in history, interacting minds or brains are abstract notions, thought experiments of the imagination; what we find in the real world are, however, only interacting living bodies or persons (2009, 468).

Criticism of social cognitivism such as these are mainly leaning on the world of everyday life and language; how we use words like “simulation” or “inference”. Accordingly, one way to respond to such criticisms is to shift the concepts of TT or ST to the *subpersonal sphere*, to say: it is not that we actively go around inferring things. The inferences or simulations that supposedly make social interactions possible are brain processes of which we are not explicitly aware. But the problem with this strategy is, as Gallagher argues, that simulation is a personal-level concept that cannot in the end be legitimately applied to sub-personal processes without making the concepts highly equivocal (2009, 469). The simulation theories, lowering these activities to the level of brain processes, ultimately apply a first-person model to form third-person “as if” or “pretend” mental states. But, as Gallagher argues, if we are not aware of the simulation, it is nonsense to say that the brain, or a part of it, is using a model in order to generate an understanding of someone else. Brains don’t form models; it is *persons* that can compare things and identify similarities between an original object and a simulation. Mirror neurons, as Gallagher argues, are equally activated by one’s own or by another’s movement; they are neutral with respect to who the agent is. They cannot, in consequence, simulate or pretend *as if* my intentions are your intentions.

Thus, there are some serious problems with a sub-personal account of simulation. We see that there is a tendency in present social cognitive science to rely on brain mechanisms such as mirror neurons or other special modules to explain social cognition (2009, 469). However, it is not at all clear whether such information more often than not only seem to presuppose the realm of the social rather than explain or illuminate it. In consequence, phenomena such as cooperation and group action end up being interpreted as an inferential or projective process encapsulated in our brains, and that illuminate *one section only* of the whole circle of organism-

environment interaction. What is missing in such theories is that the whole of the social interaction itself, as a larger structured and structuring process, in turn influences the brain functions (2009, 469). The brain, and the mind, is embedded in a context of embodied and meaningful interactions that instantiate larger patterns connecting many different levels of different processes.

Another part of the enactivists' critique concerns the problem of shared intentions. Intentions seem to be highly complex and articulate, and seem to require an advanced adult form of activity. The complex interrelation of intentions, or the forming of a we-intention, in group activities seem to depend on a high-level mental capacity. If capacity for articulated and sophisticated intentions needs to be in place for there to be joint action and cooperation, then how do persons that are not in possession of such capacities engage in such activities? Critics of the standard theories point to preverbal children, or children with autism, who struggle with the act of inferring (2014, 1). How can children form groups and engage in cooperation if this requires complex we-intentions or interlocking plans?

Thus, for the enactivist thinkers, the standard theories fail to fully acknowledge the complexity of situated and embodied personhood; history and time are important to understanding this embodiment. Valentina Fantasia et al. and Thomas Fuchs et al. call this the interactional dynamics of social reality, which points to the participants' relational and personal history and experience. Fuchs and De Jaegher criticises the current theories of social cognition of being mainly based on a representationalist view that obscures more than it reveals about actual social life (2009, 465). To the representationalists, the question is how we are able to predict and explain others' behaviour through representing their mental states. Fuchs et al. remarks that this approach tends to favour a rather narrow third-person paradigm of social cognition, where the mind constitutes a passive observer of the behaviour of others. It is only able to do so by means of an inferential simulative or projective process occurring in the individual brain (2009, 465). For the enactivists, this obliterates the actual event of the social itself, namely interaction and how interaction actually occurs between actual embodied human beings.

Several studies of the so-called "social brain", advocate that the brain is inherently socially structured through evolution. Fuchs and De Jaegher makes a fundamental critique of the "social brain" theory and similar approaches, which also include the third-person paradigm of social cognition as a passive observation of others' behaviour. Using the inner modelling process to account for social communication, these theories, according to Fuchs et al., claim that persons never actually interact with each other, but always deal with internal models or

simulations of one another's actions (2009, 466). This critique goes quite far, and Fuchs and De Jaegher fully admit that it is a caricature, but they maintain that it reveals the fundamental problem. By looking at how the brain is inevitably connected to social reality, these theories presuppose rather than explain the phenomenon of interconnectedness or interactive relation itself. Social interaction is itself taken for granted, according to Fuchs and De Jaegher (2009, 466). These standard models aim to predict other people's behaviours, but most often, when engaging in group activity, people are not consciously trying to predict anything. The group members will simply be busy engaging in conversation or in cooking a meal or dancing together. Indeed, being too conscious about one's own actions might actively obstruct the interactive process.

Imagine a couple performing a tango together; if they stop being completely immersed in the drilled activity of movements, and instead start trying to predict or simulate what is going on in the mind of the other person, they risk falling out of the rhythm. Prediction or simulation may be done under special circumstances, for instance when designing a house or drawing a face. Also, both an "explanation" and a "prediction" are specific activities that we do to achieve certain things; but these activities are not equal to the activity of interacting. The representationalist approach, enquires into how simulation, mind-reading, predicting, inferring etc. can help us reach out beyond the internal and hidden mind to the outside world and into the intentions of others. How can individual intentions align or come to share some common knowledge? However, these theories do not focus on the actual interaction taking place (2009, 466).

In the above, I have discussed the enactivists' criticism of the social cognitivist and representationalist theories. I will now move on to look more closely at the enactive approach to groups and joint activity. The conceptual language of enactment, as we saw, is that of embodiment and interaction. This approach has built on recent studies of children interacting with caregivers, including children with autism, looking at how and in what way they are able to take part in cooperative tasks even at a very early stage of infancy. As we shall see, autistic children often do not possess an explicit understanding of each partner's role in a group activity, but they are nevertheless able to partake in cooperation and therefore in group activities.

The theory of enaction is defined by Fantasia as a "non-reductive naturalistic approach that proposes a deep continuity between living and cognitive processes" (2014, 4). For Fantasia it is understood as a scientific program, exploring several levels of the life-mind continuum, exploring concepts such as autonomy, emergence, experience, and participatory sense-making (2014, 4). Enactive theory make the intersubjective process primary. This means that, for the

enactivists, consciousness and states of mind are not just inherent properties inside the private and hidden domain but viewed as emerging dialogically through the dialogical process of interaction and relational communication. The mind, in other words, is seen not as inhering in the individual, but as existing dynamically in the relationship between human beings, nature, history and their immediate surroundings (2014, 4).

Consciousness, according to this view, is not separable from the world and its environment, but constitutes a continuous dialectical “project” partaking in reality. Thus understood, the mind is formed through the shared intersubjective process of history, and consequently, the subject is inseparable from the world (2014, 4). This understanding of mind is connected to the concept of constitutive and interactive *autonomy*, defined as a network of dynamic processes, e.g., metabolic, immune, neural, sensorimotor, that continuously fuel an identity, a sense of self, under precarious conditions (2014, 4).

According to this scheme, consciousness is always situated in a world that is significant to it and which it participates in through communicative interaction. This interactive “world” is not a static pre-given space but a dialectical process that *enacts*, that is, it must be shaped as part of its embodied activity. On the one hand, the mind is constituted by the activity of the intersubjective process, and dependent on it, but on the other hand, the mind itself, in turn shapes this world through its enactive interactions. Cognition, therefore, is in practice always *embodied*. This means that a cognitzer’s activity depends non-trivially on the body (2014, 4). As Fantasia writes:

The body is more than just anatomical or physiological structures and sensorimotor strategies; it *is* the precarious combination of various interrelated self-sustaining identities (organic, cognitive, social), each interacting with the world in terms of the consequences for its own viability (2014, 4).

According to Fantasia then, the mind cannot in any straightforward way be separated from the intersubjective and embodied communicative activity, because it is dependent on it ontologically. This activity is embodied, it is not just abstract and conceptual; it must be expressed through our bodily movements, our speaking and doing.

Another important concept, *participatory sense-making*, encapsulates the relation between emergent autonomous agents and the world of significance that they together embody and enact (2014, 4). Since sense-making is embodied and enacted, it is also invariably part of a world of feelings and moods. Life through the body implies *valued* or *concerned* acting and

interacting, and this means that cognition needs affect in order to make sense of things (2014, 5). Enactive cognition relies on the sense-makers' inherently meaningful perspective on- and interactions with the world (2014, 5). From this perspective, cooperation and group activity take place in the broad and encompassing interaction, that is the meaningful engagement between subjects in which three aspects are crucial: engagement, meaning, and subject (2014, 5). Meaning and subjectivity, if we follow Fantasia, have to do with the way living (cognizing) systems are always to some degree meaningfully engaged with the surrounding environment, and which, even though dependent on the larger patterns of interactions, are fragiley self-organizing and self-maintaining (2014, 5).

The enactive approach highlights the study of *social interaction processes* as phenomena which includes different dimensions of verbal and non-verbal behaviour. This includes considerations on historical contexts, numbers of participants and technological mediation (2014, 5). This can be illustrated by looking at the joint activity of conversation. When people engage in this joint activity, the participants can experience that the conversation takes on a life of its own and social interaction becomes an experience of flow and concordance. When we, in everyday language, say that we are "being in sync with someone" or that we are "of one mind", we are in effect allowing our minds to be formed by a particular form of intense interaction process.

As Fantasia states, "engagement is the fluctuating feelings of connectedness with one another, including that of being in the flow of an interaction" (2014, 5). This happens (or can happen) because we are able to immediately recognize the body-movements, the facial expressions, and the gestures of others. The intentions of the others are not hidden, they are expressed and are gradually revealed in these embodied expressions. The intention is evident, and one speaker is able to "follow" the other and respond to the enacted "style" and "character" of the intention. It is engagement that can make this possible without even the participants being fully aware of it. Engagement, thus, reveals how social interactions are sustained by processes of embodied coordination, including rhythmical oscillations between alienation and reconciliation (2014, 5).

According to the enactive approach, there is no need for initial fully intended and rational future-directed plans, that are subsequently meshed with the plans of other group members, in order for persons to cooperate and act together. There is not even need for a fully formed intention by the participants. The embodied activity that initiate the action can be mainly spontaneous and unconscious. Participating in the interaction can change or affect the participants' goals and probably will do so in most cases of group activity. The participatory

sense-making in group activity is in many cases a scene of change, transformation, and improvisation.

This means that our intentions must be understood in a different way compared to that of the theories by Gilbert, Searle, and Bratman. For the enactivists, intentions and propositional attitudes in general cannot be understood in isolation from the intersubjective process that shapes them; intentions are truly generated and transformed interactionally. The subjective and dialectical communication between participants in joint activities is ontologically primary.

This fundamental communicative interplay opens new domains of sense-making that the mind would not have produced on its own. While mentalist approaches treat intentions as mainly hidden and only shareable by high-level cognitive mechanisms, the enactivists argue that understanding others does not necessitate inferring/simulating a special set of conditions of “I know that you know that we J etc.” We are from the very start intersubjective and relational in our bodily movements and expressions, even before we start to express ourselves through words. The intentions of the group emerge through embodied interactions that are meaningful in the given situation and in the interactional history (2014, 6).

Intentions are always tied, in some way, to a sequence of events, of a dialectical sequence unfolding through time. This structure, this history, need not be fully understood in order for it to be effective in shaping the goals of communities which in turn shape the goals of persons. There are, in consequence, different forms and aspects of cooperation: embodied in time, in space, in topic, imitative or complementary, and the fact that we are in interaction guarantees that some basic cooperative layer is present (2014, 6). Therefore, we can say that every time we interact, we cooperate, in a basic sense. Furthermore, as participatory sense-making always involves affect, we interact mainly through feelings which in turn are expressed through our bodies. Participants in groups have an embodied experience of what it means to be in that group, which must be studied if we are to give a proper account of the shared intentions of the phenomena at hand. Imagine someone coming together to start a journal. Let us say that this journal is created to allow expression of a political vision through articles, poetry, and visual aesthetics. In this case, the responsible editors act together as a group, but to understand their actions one must recognize that the activity is inseparable from an intersubjective process of embodied interaction. To understand what the group of editors is doing, one need more than analysing the intentions of individual editors. One need to look at how their vision is embodied through a certain style, a certain way of life. The embodied and artistic expressions expressed in the activity is part of what give rise to the plans and intentions of the group. The joint action is dialectically connected to enacted meanings. What is communicated through the shared

action of creating the journal is conversation, an intersubjective process of meaning-making. In this way, anticipating, collectively shaping a product, the group activity expresses the emergent vision of what is going on in a particular community. Take away the artistic, the enactive, and the embodied level, and you end up having difficulty making sense of the shared intention and action of the editors at all.

The intentions of a group should not only be sought as something that lie “behind” the communicative action as its “cause”; the intention manifests itself in the speaker’s actions and speech as well as in his or her body language through space and time. It is not that human beings are not able to hide their intentions, but that the ability itself of hiding their actions would not be an issue if not intentions usually were openly manifested through our embodied action and speech.

3.2 Enactive theory on joint attention and language acquisition

In the following, I will continue the discussion of whether participatory sense-making can make better sense of how we act, think, and cooperate in groups. One benefit of the enactive approach is that it takes into account the intersubjective and historical processes that underlie and support our normal everyday sociality. Another point is the focus on language acquisition in young children. The enactive approach perceives the consciousness of children as being shaped through engaging in actions and practices together with others. It is the communal participation in practices, in itself, that enables children to acquire language and concepts. By engaging in embodied practices of cooperation and shared attention, children are able to acquire the concepts and words that will later enable them to partake in the whole realm of language. Cooperating is thus possible even for those, such as young children, who do not possess a robust conceptual capacity to “read” other’s intentions or plans (2014, 6).

Let us look more closely at the enactive approach to cooperation and shared attention. In recent years, the interaction between infant and mother/caregiver, has been a focus of research, particularly regarding how infants are able to acquire concepts and language. In the paper “How Wittgenstein can help us understand autism better” (2009, 2009) Ljiljana Radenovic and James Connelly present a perspective that may illuminate the significance of joint attention. Young children, according to Radenovic et al., are not able to determine necessary and sufficient features or the most frequent relevant features, when faced with the objects of the world (2009, 5). Their Wittgensteinian approach, which resembles that of the enactivists, argue that a child’s ability to see similarities or patterns, to notice regularities, or discriminate

between objects and events, is revealed by their ability to engage in shared activity or joint attention. If the child can use a word or concept while partaking in joint attention and activity, it has learned the concept or word. The child's acquisition of the concept is, therefore, identical to his or her ability to classify things into proper categories (2009, 5).

The ability to perceive similarities is part of what it means for the child to understand the concept; to be able to classify objects into proper categories is part of what it means to possess the concept. Radenovic argues, following Wittgenstein, that one should consider the grammatical or normative connection between perceiving similarities and understanding a concept, or grasping a concept and applying it, as constituted within the context of rule-governed practices. These practices themselves are typically situated within a broader social and institutional background (2009, 5). In this sense, it is the social and communal that is the irreducible whole that makes sense of the skills and habits of a particular community and its members. This approach is in line with the enactivists, as we find here a form of holism where the whole of the shared practice cannot be reduced to its individual members, but rather remains the context that makes sense of our practices and our language.

Given this conception we can ask what are the routines whereby the child begins to master a normative practice? What enables the child to participate in such practices and what exactly in these practices enables the child to develop cognitive and linguistic abilities? Radenovic explains how children with autism often have deficits in abstract reasoning which are displayed as atypical learning strategies in the acquisition of concepts and categorization. These children show an inability to generalize, that is to recognize the larger patterns, the wholistic landscape within which we make sense of particular words and concepts.

Like Radenovic's Wittgensteinian approach, the enactivists see the relationship between perception of similarities, possession of a concept, and our ability to apply it in a particular practice, as what must be understood as how the child learns to *engage in joint activities*. The enactivist looks at the ongoing interaction between the child, the world, and other people. The social learning takes place in such interactions. The child's ability to learn through social interactions is probably present from birth, but accelerates as the first "joint attentional scenes" emerge. With "joint attentional scenes" the child starts being able to make use of the communicative signals of a caregiver, which means that the child is able to follow the gaze of the caregiver and determine the object(s) to which the caregiver intends her to attend. For autistic children, on the other hand, literature suggests that the participation in joint attention emerges very late in their development (2009, 6). Consequently, it is likely that the considerable delay in joint attention is the main culprit in the delay of language acquisition and the impaired

development of social cognition in these children. This same lack of joint attention is then, plausibly, the main cause of impaired concept acquisition and categorization abilities in these children.

Studies of social interactions between infants and caretakers have shown that participants can coordinate their movements and utterances without any conscious or intentional effort. This is the case when a baby directs its attention towards the same things as the mother. Since infants cannot stay alive alone, they need others to help them with nourishing, shelter, hygiene, and social interaction. Fantasia et al. refers to research showing that it is to be expected that infants contribute actively to this caring, because they are themselves sense-makers, generating and maintaining their own living identity, and already their social identity (2014, 6).

Following this research, cooperation in infancy is defined as the joint management of objects, actions, or ideas to fulfil a purpose that two interactors share. Both the infant and the caregiver contribute to the joint attention that emerges through the early communicative interactions. The infant without autism spontaneously partake in communicative interactions involving giving and receiving objects before 1 year of age. Recent observations have demonstrated how, from very early in life, infants adjust and facilitate actions directed towards them. This is especially seen in daily routines such as when the caregiver picks them up, change their nappy, or play a social game with them. The process of cooperating enables children to build up their actions moment by moment through a sequence of relational adjustments and engagements towards a joint goal (2014, 6).

Fuchs and de Jaegher present a concept of social understanding as an ongoing, dynamic process of “mutual incorporation”. By “mutual incorporation” they mean a process in which the lived bodies of both participants extend and form a common “intercorporality”, a shared intersubjective space of embodied meaning. This is evident in the dyadic interaction in early childhood (2009, 465). In the joint attention shared by the child and the caregiver, a fundamental incorporation takes place, constituting the source of later communication. In contrast to theorists like Bratman and Searle, who in different ways both underline the importance of rational intentions and reasoning for sociality, Fuchs and De Jaegher present a non-representational, enactive, and embodied concept of intersubjectivity. Here, social understanding and action is not realised by “snapshot” activities resulting from one individual’s intentional theorising, planning, or simulating, but arises in the moment-to-moment interaction between two embodied beings.

This kind of interaction involves coordination but does not imply perfect synchronisation. Studies show that non-autistic infants from 3 months of age prefer slight modulations (time-delays) and imperfect contingency in responses (Gallagher 2010, 115). In these studies, the oscillation between attunement, loss of attunement, and re-established attunement maintain both differentiation and connection in the case of joint attention with the caregiver. The creation of meaning is a result of the intersubjective process itself, and goes beyond what each individual can bring into the process. When two people dance the tango, something dynamic is created that neither of the dancers could create on their own, and this same form of dynamic communication is already evident in the interaction between infant and caregiver. This is why interaction has a certain autonomy. Neither of the participants are able to completely control the interactive process, given that it is truly interactive. In cases where one person is totally in control of the other person, there is no interaction understood in this specific sense. Indeed, this reflects a basic “hermeneutical” aspect of interaction – interpretation of the other person, or of the world *with* and *through* the other person, always extends beyond the control of the agent, and this is the basis of learning. Accordingly, sustained and repeated interactions build “implicit relational knowledge” and improve possibilities for greater fluency, flexibility, and further successful interactions (2010, 115).

The intentions of the other person are communicated in their bodily movements, grasping, pointing, moving towards, moving away from, staying close, nodding, gazing in a certain direction, etc, all occurring in specific ways, often un-consciously. Interactions also include mutually generated intentional actions and responses that lead to chains of meaningful joint action (2010, 116). One intentional movement, that of pointing, may motivate movement on the other’s part, moving in a specific direction, reaching, grasping, moving back, handing over, taking back, and so on, creating a “common space” of meaning in movement.

Let us return to children’s interaction with their caregivers. In joint attention, which begins to emerge around 9 months of age, the child learn about the world through others (2010, 116). When the child sees another person look at something or gaze in a certain direction, he or she automatically tends to follow that person’s gaze to the object or the location. This is often the initiation of an instance of joint attention, which is important for our ability to generate meaning in participatory sense-making. How our consciousness perceive the world and what patterns and regularities it is able to identify, are to a significant degree shaped by the gazes and actions of others.

The enactivist thus suggests that the task is not to access interior propositional attitudes or fully formed intentions, but to look at the motoric and spontaneous gestures and bodily

affective movements of the child. Connected with this joint attention common to the infant and the caregiver is the fundamental “inter-action” which in time helps the child to acquire language and the ability to participate in complex forms of joint action. This interaction, according to the enactive point of view, is an intersubjective process that makes the joint attention possible.

How are we to understand this intersubjective process? Here I think it is helpful use music as an analogy. When a piece of music is performed, the emerging music is the result of the composer’s description in the music sheet as well as the performance by each musician, in time and space. Similar to music, the intersubjective process upholding and empowering all social and personal reality, constitutes a complex rhythmic flow. Take away the intersubjective process from our account of cooperation and joint action, and about the result will be philosophical abstractions only applicable to rational disembodied minds, not the incarnated, dependent, and finite persons that we actually are.

Language and concept-acquisition can only develop and exist because we already are able to participate in the intersubjective process. The mind is through and through relational and our joint activities can be almost unconscious or dreamlike, and yet highly cooperative, as when a skilled couple engage in a dance. We are able to interact with one another and the reality as a whole, in a myriad of ways, but first and foremost in an improvising way, and we are never able to have more than limited control of the outcome of a situation. We are social because we participate in an ontological “pattern” that makes meaning and sense possible. The enactivists do not, however, address the nature of this pattern, other than pointing to the interactive process itself. In the concluding chapter we shall address the question of the background patterns of the social world, but for now we continue to focus on the enactive theorists.

In support of the enactive approach to sociality, Gallagher introduces the following phenomenological example:

I find myself in a pasture looking at some cattle that start to move toward me. As the individual cow moves it seems to be checking that its fellow cows are coming along. It seems that the cows themselves are involved in a joint attentional situation. My own response is not to try to get into their minds, or to work out some set of propositional attitudes that I attribute to them. Rather, my understanding of their joint attention on a particular object (me, in this case) and my own attempt to make sense out of what they are doing immediately translates into movement on my part. Furthermore, if I try to discern whether they are going to change direction before they reach me, I don’t try to mindread their mental states, I try to see it in their movements—or in the movement of the herd as a whole—and this in reference to the shape of the field and the

possibilities for their moving this way or that. My attention to what and where they are attending sets my feet in motion (2010, 118).

If we consider a variation of this setting, Gallagher argues, we can see that the attention that the “I” share with other humans always involves something like this. If, for example, two friends happen to be in the pasture looking at a cow herd coming towards them, if they catch each other’s eye, and if one grabs the other’s arm and yells that they should run, is there anything more to the coordination that we have to explain? Do the one friend need to start “reading” the intentions, beliefs, and desires of the other in order to cooperate in the situation? Do they need to employ a theory to understand why someone grabs another person’s arm? Do the one mind need to simulate the other’s thoughts on the situation? Here the kind of knowledge the two friends need is a very practical kind that is based on their ongoing participatory sense-making and spontaneous interaction.

Notice that in this example we are engaged in both social cognition, that friends are able to make sense of the others gestures and body-movements, *and* participatory sense-making (together they are jointly attending to and making sense of the herd’s behaviour) (2010, 118). This is then an example of spontaneous social interaction. Bodily action, facial expression, gaze direction, tone of voice, bodily posture, and gesture, communicate the specific meanings of the group as they participate in activity (2010, 119).

Joint attention and joint action are thus deeply connected; the former is maintained through bodily and communicative movements (action on some scale), and leads directly to a complex joint action that is guided by the continuing movements, postures, gestures, etc. of all participants in a community. A common intentionality emerges among the individuals as they enter into the interaction involved in a specific task. This interaction is carried out primarily in spatially and temporally coordinated movements that include action-oriented positioning and posture, and communicative gesture.

To understand this process, we need to study the entire group as the unit of analysis. Their shared understanding emerges from a set of embodied movements and actions in the specific context of what they are doing, and it is irreducible to any set of mental states in one individual or even the collection of mental states in all of them. The action, and the meaning of the action, transcends any one individual; it is generated in the interaction required for the outcome. The joint action is a distributed and collective phenomenon (2010, 120).

Thus, both social cognition and participatory sense-making, in cases of joint attention and joint action, depend primarily on *movement* – bodily movement and posture, taking up a

relevant position *vis a vis* the other, gestures, facial expressions, vocal intonations, communicative actions, and so forth. Evidence can be found for this interactive approach in an extensive array of developmental studies, in phenomenology, and in empirical observations from the field (2010, 120).

Returning to the case of young children, Fantasia et al. argue that even children that perform “poorly” in interactive communication with the caregiver still show a capacity for joint attention. Indeed, some children with autistic spectrum conditions (ASC) do manage the cooperative tasks (2014, 6). Based on this, studies of the verbal production of children with autism have been performed, not focusing primarily on their deficits, but trying to understand these children’s spontaneous interactional behaviour. Fantasia argues that conversation analysis studies support that the autistic child’s repetition of utterances with no apparent relation to prior talk from other speakers, are in fact responsive moves. Repetition of available utterances assists these children in remaining active in the conversation despite their difficulty with improving a newly designed turn (2014, 6).

Observing what children are actually able to do reveals capacities for cooperation that cannot emerge in pre-defined tasks. One investigation studied classroom interactions between two autistic children and their tutors. The teachers posed questions and the children were asked to answer, using answer-cards. As the session went on, each of the children acted by tapping the answer-cards. This action might at first sight seem meaningless, but the children in fact tapped on the cards just before they started answering, and sometimes continuing into their answering. The children were thus able to find means to signal their ongoing engagement when the timing of their verbal production was delayed. By use of these signals, they cooperated to maintaining the interactive process (2014, 7).

Such studies indicate that the use of non-verbal and non-vocal resources can be sufficient for building a co-participatory model of how children and teachers work together (2014, 7). They are able to do this because participating in the patterns of interactions does not require more complex social skills or intentions. The children with autism are able to partake in these exchanges by “growing” into them, that is, the everyday cooperative interaction teaches them to be active partners and develops their participatory sense-making (2014, 7).

Following these insights, we can say that actions need to be understood in their interactional context and in their significance for all participants. With that perspective we can more fully appreciate the emergence of cooperation also in the interactions of- and with people with autism (2014, 7). This is the merits of the enactive approach, opening up a conceptual schema for making sense of a wider range of cooperative interactions, not only those in which

interactors explicitly agree upon and set rules for the conjoining of intentions and plans. There is no reason to reject the stricter concepts of group intention and joint action, which certainly provide important insight regarding particular scenarios where participants must make an effort to make sense of each other's intentions, and where a shared goal must be established beforehand. However, the enactivist approach shows that on a fundamental level, group activity depends on broader and more encompassing patterns of interaction, and that interaction allows for a multi-layered process that may take different forms. Intentions-in-action are shaped and transformed through the ongoing interaction and shared sense-making of groups and communities acting together. For collective intentionality to come about, a form of participating in each other's sense-making needs to already, to some degree, be in place. The joint action performed by adults is an already fully developed one, but it cannot stand as a universal paradigm for all cases of cooperation and communal activity; observing infants and children with autism helps us appreciate more basic ways in which such activity develop (2014, 8).

3.3 Strengths and weaknesses of enactive theory

There is certainly value in adding major thoughts within the enactive theory to the existing accounts of shared intention and cooperative activity. However, to the degree that enactive theory tries to present an alternative to the standard theories, I do not think it hits the mark. Indeed, I would argue that insofar as theorists like Fantasia and Gallagher criticize the standard theories, those of Gilbert, Bratman etc., for being too "mentalistic" by making complex adult intention fundamental to joint action, they are attacking a strawman. Let us consider Bratman's theory, perhaps the most developed standard theory of joint action. The idea that his intentions or plans are too conceptually complex and rationalistic, that it would not allow for cases of sociality that are more spontaneous, triggered by enactment and gestures, is simply untrue. Bratman's planning theory is meant to be relatively neutral with regards to particular theories in social cognition or the philosophy of mind in general (although he admits that he adheres to a functionalist philosophy of mind⁷). Bratman does not construe human beings as disembodied rationalists, constantly planning, eschewing spontaneity, and rigidly acting according to prior plans (2014, 24). He agrees that there is a danger of arriving at a one-sided portrayal of human agency, but he reminds us that our planning capacity, and our intentions in general, are always embedded in a complex psychic economy that also involves abilities to characterize one's plans

⁷ Bratman writes "My approach to intention is, broadly speaking, within the functionalist philosophy of mind" (1999, 9).

in schematic and conceptually open ways. Our group action can be spontaneous, and our plans can be only a rough sketch, whose contours and details can be filled in at a later stage in the collective action. We need to do justice both to the centrality of planning and of agent rationality in the constitution of fundamental forms of social organization, and to our important capacities for conceptual openness, spontaneity, and flexibility (2014, 14). But the attempt, made by the enactivists, to move completely away from intentions, and from the centrality of the mental in the constitution of joint action and interaction, risks, I think, leading to a form of behaviourism. By “behaviourism” here I mean the general methodological and philosophical presupposition that there is no principle difference between the behaviour of human beings and any other process in nature, that the former can be accounted for in the same way as the latter, by laws relating to physical events, and that the introduction of such notions as “intention”, “purpose”, and “mind” can only serve to obscure and confuse⁸. The enactivists obviously want to discuss meaning and shared communication; indeed, it is one of their central points that embodiment is expressive of meaning, and that we should understand mind and intentions as emerging from a shared communal and embodied history. Still, the radical eschewal of “intention” in their model of sociality, place them in danger of proposing a behaviourist construal where participation and cooperation are no longer essentially connected to the mind. And their attack on the standard understandings of sociality, whether reductive or non-reductive, seems to be more of a strawman-argument, than an actual confrontation towards these theories.

We can however value the enactivists’ “embodied” approach, and their interpretation of the research on the children-caregiver interaction, without accepting the enactivist theory as such. If we follow the enactive perspective, group cooperation and shared intention partake in the larger intersubjective patterns and rhythms of embodied interaction. The background “patterns” or “rhythms” seem indeed to be insufficiently discussed among the standard theorists. But how are we, for example, to understand small-scale “modest” forms of sociality in light of larger historical and dialectical rhythms of the large-scale social world? Is there a theory which can better account for the insights of both the standard and the alternative enactive theories?

Bratman’s continuity thesis proposes the idea of interrelated individual intentions as the starting point for understanding cooperation as such. In one sense such intentions and meshing subplans are central to cooperation and joint activity, but in another sense, they give

⁸ This is very general definition of behaviourism is proposed by Charles Taylor in *The Explanation of Behaviour* (2021, 4).

expression to something more primordial than the plans, themselves. People may gather together to act and speak without having a clear plan of what will unfold while doing so, something Bratman also points out (1999, 29). Shared goals or plans may emerge during the course of an interaction, and so participants can fall into a rhythm or pattern of cooperation without having a previous thought or clear intention.

Not enough attention has been focused on the phenomena of *falling into* certain shared activities or patterns of habits: one can fall in love, for example, and thus express a certain pattern of habits or moods for a long time. One can “fall back” on old habits or old ways of speaking; one can retreat back to old ways of thinking, and so on. These examples show that there seems to be some underlying patterns that make sense of our social habits and actions. For instance, two friends meet each other unintendedly in the street and immediately start a conversation. Each adjusts movements and postures; maybe first the gesture of greeting, and then, without saying a word, they *fall into* a pattern of shared activity, a conversation. They did not, strictly speaking, choose to engage in the conversation as it started before they had time to make a rational choice. Is there a shared intention present here? One could say yes if by intention we refer to a very general sense of the implicit directedness of consciousness. But if the term is used in a more explicit sense corresponding to the phrase “I intend to have a conversation with you”, it would be more correct to say that the intention/the common goal emerges as a result of the two friends first *falling into* the pattern of interaction. In this case, the *habits*, the rhythms, and patterns which lie potentially in the body of the persons, seem to be more primary than the intention, if the latter is understood in a strict sense of “I intend to have a conversation with you.” In the next two chapters we will explore another approach, or set of similar approaches, to group action and cooperation, that might better integrate the insights of both the standard theories and those of the enactivists. We are ready to take on the classical phenomenological tradition as it is presented and defended by the philosopher Dermot Moran.

Chapter 4: The phenomenological approach

4.1 Introducing classical phenomenology

The classical phenomenological approaches to sociality, on acting and thinking together, focus on the larger underlying patterns that constitute the social world. When looking at small-scale cases of “modest” sociality, the phenomenologist would argue that such modest theories cannot make complete sense of small-scale actions. The phenomenologists have in common a strong emphasis on the way human action partakes in history and in particular communities and cultures. For them, the individual develops out of a group and the process of individuation is essentially intertwined with that of socialization (Moran 2021, 10). Thus, for many phenomenologists, the larger historical pattern is the key to understanding collective action.

There has been an ongoing debate regarding collective agency within the phenomenological tradition, almost completely ignored by current Anglophone discussions. For classical phenomenology, as it is presented by Moran in “The phenomenology of joint agency” (2021), there are many different degrees of joint or shared agency with different levels of agential involvement, different degrees of responsibility, and so on. In consequence, Moran is sceptical of approaches that try to provide an exhaustive explanatory scheme, and favour approaches that is conceptually open to change registers and conceptual vocabularies depending on the social group at issue. There is no one essential way of describing social groups. The groups may vary in features such as size, hierarchy, distribution of labour among members, types of interaction among agents, transience or permanence of the group, dependence on institutions, which are all essential to the different concepts needed to understand them (2021, 7). Depending on the groups and on the practices, they each have their own attentional focus, anticipations, visions, and fulfilments, within defined temporal frames. Even the frames of temporality may be different in crucial ways. Each type of joint agency has its own peculiar and very complex phenomenology that needs to be investigated using the phenomenological method (2021, 7).

Now, phenomenologists are not alone in recognizing the plurality of social forms; Margaret Gilbert’s discussion of groups does, within the standard theories, recognize the many differences: “Clubs, trade unions, and army units are likely to have a set of explicit rules of procedure and explicit goals. Families are less likely to have such rules and goals” (2006, 94). She acknowledges that larger groups may have members who are unknown to each other, but her approach is still to begin with small groups, two people going for a walk, painting a house

and so on. Gilbert, as well as Searle, Bratman and Tuomela, hardly discuss the alternative tradition, and they all tend to limit themselves to small-scale cases where their theories are difficult to criticise. Phenomenologists, on the other hand, prefer to understand how smaller groups belong to a larger social world (Moran 2021, 18).

Shared social groups and social roles need a richer and more historically oriented philosophical vocabulary in order to account for the way joint action in different contexts can be spontaneous or premeditated, temporary or permanent, or how it may simply befall one. When recognizing this, we also need to acknowledge that the intrinsic forms of cooperative and collective action involved also vary and can range from the fully implicit (doing things with one's family and friends) to the highly explicit and articulated (constructing a philosophical article, creating a legal contract etc.) (2021, 6-7). Within groups we find radically different distributions of engagement, responsibility, recognition, and reciprocity. In consequence, when we want to make sense of an action, we have to make sense of the story or set of stories providing the context that makes the action intelligible.

Moran presents a family of views within the classical phenomenological tradition and connects them to the discussions regarding joint action and social ontology. He asks: "does the phenomenological point of view, does joint or shared agency need a conscious sense of shared agency?" (2021, 1). For phenomenology, understanding habits and habitual patterns as constituted by *a shared horizon of meaning*, is crucial. He points out the ubiquity of cases where human beings partake in different groups activities without reflecting about what they do: "there are many processes where we seem to just go along with the group without conscious intent" (2021, 1). What makes this possible? For phenomenology, something more than planning, beliefs, motivations, or habituated dispositions is the source of our fundamental habits. Building on the classic phenomenological accounts by Edmund Husserl, Alfred Schutz, Martin Heidegger, Jean-Paul Sartre, and Berger & Luckmann, Moran emphasizes the thick horizon of *the life-world* as a fundamental condition for intentional shared agency. All the different forms of joint agency, of joint action and attention, are embedded in a pre-predicative tacit knowledge in this overall life-world (2021, 1).

Husserl, Heidegger, and Schutz have different accounts of the term "life-world", but their views have in common an "anonymous," pre-predicative kind of group participation. This anonymous shared social "space" represents a new concept of agency, one that is nevertheless neglected in the current literature on philosophy of action (2021, 1). In phenomenology, human beings live their lives embedded in an encompassing, largely taken-for-granted, intrinsically social, temporal, and historical world (2021, 2).

This intrinsically social and historical dimension of human nature implies that the individual lives, self-conceptions, and agencies, are pervaded and saturated by others. A major question, addressed by Husserl, Heidegger, Sartre, and others, is how the individual consciousness intermeshes with the collective, intersubjective domain. We remember that the enactivists discussed an underlying “intersubjective process” constituting social interaction although the enactivists’ meaning of intersubjectivity remained unclear. The concept of intersubjectivity within the classical phenomenological tradition seems to be more helpful for the understanding of the underlying social structures that make human social action possible. In this tradition, intersubjectivity is connected to the notion of the life-world, the anonymous presence of others, to the whole of the historical community, in the actions of groups. Rather than following the individualist route, phenomenology try to show how the larger groups constitute the enabling condition for the individual agent. The agent may certainly rebel against the group, and it will in most cases mature and develop its own self-consciousness. However, this is only possible within the larger groups that form our languages and ways of understanding the world.

Like Bratman, phenomenology also stresses that our actions are structured by norms, but the focus is on how norms of group action are inherited and cultivated in historical communities. The social norms which structure the condition for our normal way of being and living together with others, are largely passively inherited. When we as children first inherit them, we do so without acknowledgement, or even awareness, from our largely anonymous “predecessors”, as well as from our “significant others”, i.e., family, friends, those involved in our care and nurture. We might later in life discover the cultural and historical origins of our ways of life, our manner of speaking and perhaps we find that a certain concept or ideal that is important in our lives, “liberation” or “autonomy” for example, originates from the Enlightenment and the philosophy of Emmanuel Kant. These social norms and ideals are often inherited during the first stages of acculturation, during childhood.

Several theorists who have concentrated on the concept of “construction”, the notion that we construct, through collective intentionality or joint commitment our shared social reality, have left out the important notion of “receiving”. We may say that shared language, communal norms, values, and what may be called “knowledge” in a broad sense, are received as a gift. These inherited abilities, language, and social norms may be good or bad, true or false, but it is important to understand the background norms and structures that make human beings capable of acting and attending together in communal practices. By overemphasizing the undeniable notion that we participate in the construction of social reality, we risk overlooking

how social groups nevertheless depend on something more primordial, which is received and inherited.

The analytic approaches, for example those of Bratman (2014) and Gilbert (2006), assume that concepts such as shared intention, joint commitment, and planning involved in individual agency or small groups, can largely be mapped, and conceptually understood without addressing the shared meanings that such small-scale groups constantly receive from the larger social domain. In the standard theories, as we noted, one need not look at the larger cases in order to understand the smaller ones. Rather, the analyses of small-scale cooperative activities, are used to build larger models used to better understand how the larger social structures function.

We remember that for Bratman, the idea of individual planning agency was sufficiently rich in its conceptual and metaphysical vocabulary to be applied to cases of larger historical, economic, and political structures and institutions (2014, 8). Similarly, Gilbert claimed that “starting small” is regarded the best approach to phenomena of the social, because the “ground-up” explanation is regarded to be more adequate than the “top-down”. In contrast to these and similar assumptions, phenomenology understands group dynamics in social practices to require unique concepts not found at the individual level, which, I would claim, make better sense of group action than the notion of “we-intentions” or “we-modes” (2014, 2). To fully understand actions, beliefs, habits, dispositions, as well as the vocabulary we use when thinking rationally, one cannot remain at the local individual level. Such actions must be interpreted in light of the larger social practices and the communities in which they take place. Thus, in our everyday lives, we are involved not solely with immediate, significant others in our present zone of one-to-one, face-to-face relations (dinner with family, coffee with friends etc.), but crucially also with anonymized, unknown others, those of my generation, of citizens in my country, and relations belonging to history (2014, 2).

Imagine the following scenario: in the event of an upcoming election two politicians representing two popular parties are engaging in a debate. In virtue of representing the party, the speech-actions of the politicians are shaped by the formative social forces of the anonymous people that together constitute the respectable parties. These anonymous others are not present in the room, but their otherness, in virtue of together constituting the party, is nevertheless what makes sense of the conversation between the politicians. If a viewer, who watch the debate on television, wants to understand the full meaning of what is being said, he has to understand, to some degree at least, that the words used also function as representative acts, representing the interests of people that are not there, namely the members (and even the voters) of the party.

Take away all knowledge of the anonymous others and the intentions and purposes of the actions become significantly less intelligible.

This seem obvious, one might think, when considering the motivations and rational deliberations of politicians in a debate. Politicians engaging in a debate is obviously a fairly complex case, not at all comparable to the decidedly “modest” cases we find in Bratman’s philosophy (although he does use a political debate as a case-study in one of his works⁹). Let us move to a “normal” conversation between friends. Can we in such a case just focus on the “modest” part of the interaction? I agree with Moran and the phenomenologists that the holistic approach is just as important for a philosophical appreciation of conversations among friends. Synonymous, unmentioned groups and institutions are to some degree present even in the simplest of conversations, and to make sense of the utterances of a person, understood as speech which is also a form of action, we need to interpret it in light of larger social contexts, larger social entities that cannot be adequately described using only concepts of individual actions and persons.

Imagine I’m visiting a store trying to have a conversation in French with a native Frenchman near the city square of Caen. What he does with his words, that is, why he uses them and what they mean, will be difficult to understand if I’m not able to understand French. I might try to interpret his body language, and try to discern what is the most likely meaning and purpose of his words, but to do so I still have to possess a minimal knowledge of the situation he is in. But to understand what situation he is in, I need to understand, roughly, his relationship with his environment, with the people around him, and what professional or personal role he place in the larger network of social reality. Note that this rough sketch might at first be stereotypical, a lot of what I deduct from his clothing style, his manners, and his role in the social and historical community that he is part of, might be an exaggeration or even an illusion. But without me trying to make a rough stereotypical sketch of these larger social and historical patterns that make sense of his body-language, of his movements, his gestures and speech-acts, he will remain wholly unintelligible. Assuming the presence of the anonymous others, of his relationship to the historical groups that gave him his education, language, job, even when I have a partly distorted view of them, is nonetheless necessary to make sense of his individual speech and action. Classical phenomenology developed several strong accounts of the intentional subject’s involvement with unknown other subjects, with the “other”, alterity. This

⁹ See the Bratman’s discussion of the “Mondal case” in *Intentions, Plans, and Practical Reason* (1999, 77).

might be a way to interpret what it means that the anonymous others are present and formative, and must be accounted for, even in small-scale social settings.

In consequence, following this line of argumentation, formative social forces are irreducible phenomena that cannot be made wholly sense of by individualist concepts. In addition, these phenomena must be taken into account even on the small-scale level of personal rational thought and deliberation. The formative social forces are in a sense all-encompassing in that they leave little room for creative appropriation or modification by individuals. “Such shared, contextualizing social ‘knowledge’” writes Moran, “[is] a necessary condition for the possibility of joint action; and this social knowledge, in turn, is embedded in our ‘being-in-the-world’ (In-der-Welt-sein) with others (Mitsein)” (2021, 2-3). Thus, the cases I have made were by no accident cases of conversations, because the reality that social knowledge always presupposes larger social wholes, is supremely evident when looking at language. The first language the child acquires is absorbed from others in its surroundings; it is not constructed or invented, although the child may mispronounce or misconstrue (2021, 3).

4.2 The Phenomenology of Intersubjectivity

At this point we will move on to look more closely at Husserl’s term *life-world* (Lebenswelt). The life-world, for phenomenology, is the collective context, the ontological background, that makes the pre-predicative, unthematized forms of embodied and embedded intentionality. It is this background that, whether of unconscious intentionality or more explicit intentionality, whether singular and collective, provides a shared social context of joint action. Moran writes: “Joint agency emerges from this presumed world of implicit normative “sense” (Sinn) that contextualizes the action and endows it with significance. This is the communal life-world” (2021, 8). Joint action can be made sense of, as actions participating in a larger collective and dynamic background. Does the embeddedness of our intentions and actions imply a kind of “cultural determinism” in the sense that all of our are thoughts and actions are predetermined by our cultural communities of language and shared meaning? I think the shared horizon of the life-world is more like the stage where our thoughts and actions participate in an unfolding historical drama. Without this background, this public and cultural stage, the thoughts, and actions would be difficult to make sense of, but it does not follow that our actions are predetermined. Our shared actions *participate* in this larger historical and cultural life-world; they are dependent on it, but our actions can also contribute in shaping and reshaping the background itself.

The communally shared, pregiven, everyday life-world that acts as the meaningful temporal “horizon” for all action, is then the context of our “embeddedness” (Einbettung) (2021, 10). For Husserl, the life-world implies that each soul also stands in community with others which are intentionally interrelated in an internal and essentially closed nexus of intersubjectivity (2021, 13).

I agree with Moran that the Husserlian concept of the life-world offers a fruitful complement to Bratman’s approach that explicate joint action wholly in terms of individual intentions, commitments by individuals, or as individuals intervening on the intentions of others (2021, 3). Reducing joint agency to “a set of diplomatically cooperating individual agencies”, risks misconstruing aspects of sociality that are essentially connected to community, language and history (2021, 3). Theories like those of Bratman are not necessarily incompatible with the more holistic phenomenological philosophy of the social. Rather, I believe the aspects of planning, cooperation, and intentional action, carefully explicated and analysed by Bratman, will make even more sense if integrated into a richer ontology of the social.

The benefit of Bratman’s account lies in the detailed analysis of how mutual responsiveness, commitment to joint activity, commitment to mutual support, and common knowledge of these commitments, are held together by our capacity to form plans. The mutually interlocking plans of Bratman may still benefit from Moran’s approach of joint action which emphasises how the participants need no explicit knowledge or commitment (2021, 21). The musicians in a joint activity of playing a jazz piece can be only half-aware of what they do, being completely caught up in the “higher dream” of the music itself. Even so, someone who wants to make sense of their actions, likely needs a thorough knowledge of music, culture, and history, in order to make his or her account of the musician’s actions intelligible. These examples support Moran’s statement that no single conceptual account can describe all forms of joint action. Rowing together, playing in an orchestra, performing surgery, engaging in a war, are structured in different ways with different forms of participation, recognition, reciprocity and horizontality, and these joint activities are quite likely in need of different conceptual frameworks to do them justice.

We have seen the pitfalls of reducing social institutions and groups to the individual intentions, attitudes, and beliefs. However, is there a risk that phenomenology can reduce the individual mind and intentional action to nothing more than functions, cogs in a collective machinery? Is it collectivist in this sense? I think not. Although we are all participating in the anonymous collective vehicle of language with its inherent language games, we still have a sense of what is uniquely our own. Speaking one’s own native language is a way of knowing

oneself, of plunging into the depths of one's consciousness, and yet is also the paradigm for participation in a larger collective activity. Language, understood in this sense, *is* exactly a larger form of group activity, a continuous conversation through history.

That every new sentence articulated by a person is already embedded in the life-world of unknown others, does not mean that the individual's utterings aren't also unique. A unique "voice" would not be recognized if it did not participate in a larger conversation that made the uniqueness of the "voice" intelligible to others.

From birth, human beings are thrown into a conversation that has already started and is running on its own rules. It is shaped and re-shaped, sometimes actively and sometimes more passively, by those participating in a linguistic dimension, a "horizon" of meanings, symbols, expressions, and ways of life. While the overall approaches of the phenomenologists differ, there is a great deal of converging ideas, for example about a deep, embodied sense of "I", the zone of "mineness," and "not-I", zones of familiarity and unfamiliarity (2021, 3). Moran, for example, mentions Heidegger's idea of "the one" or "the they" (*das Man*), where the "I" do as "others" do. Here individuals just fall into the established patterns of habits, falling in with the crowd-behaviour and "go with the flow" or "live along with" others (2021, 3). But even in these contexts, of finding oneself lost in a crowd, there is still a background sense of familiarity, of shared horizon, of mutual comprehension, in order to act jointly with others. In such contexts, our intentions may be more or less conscious, our actions more or less conscious, depending on the intensity of self-awareness. The self is formed through its commitments, habits, thoughts, and actions in the world, principally on the occasion of my contacts with "an other", that is, in a social context. Individuals are always involved in communal groups, whether they consciously know it or not, and whether they specifically agree to it or not (2021, 4).

Now, how does this help to make sense of group action? Let us imagine strangers sitting in their cars at the ferry slip waiting for the ferry. The ferry arrives and the drivers start forming a queue; each car will be directed by ferry workers in order to enter the ferry one by one (or if it is a larger ferry: two by two). This action can be described in Bratman's sense as a form of joint activity. I intend that we J (drive into the ferry), I intend that we J in accordance with and because of meshing subplans of the other drivers, the other drivers in turn intend that we J, also that we do so in accordance with and because of meshing subplans of me and the other car drivers. The intentions involved are not coerced by the other participants (in a normal case, at least, when all the car drivers are willing to cooperate) so subplans are in concord with the other subplans of the drivers. And finally, all this is common knowledge between us, the drivers. Roughly, Bratman's account of the cooperation fits. And yet, this does not seem to account for

the more or less spontaneous action of a group of strangers; how are they able to act so well in concord? It is not that Bratman's individualistic account is false, but that it leaves out many important aspects regarding the underlying structures that enable us to cooperate in such a manner in cases like this. At other places and times, the very same people would not be able to cooperate at all. Perhaps, given a different context, they would start to fight and would for instance struggle to cooperate together in traffic.

The above seemingly spontaneous action requires, according to phenomenologists, more than intentions, beliefs, plans and meshing subplans; they require a shared culture of signs, signals, symbols, and meanings that allow these strangers to live and participate in traffic in relative peace. In western Norway there is a ferry-system in place which encompasses certain essential rules that have to be followed, certain habits cultivated among the ferry workers, a certain differentiation of social roles, and a pattern of activities extending through time. These all require that there is an essentially shared horizon of patterns, inherited to us through history. Group-being is often "mindless"; the drivers need not reflect about what they are doing, they can simply participate in the shared activity, because they partake in this larger life-world. Their intentions are certainly also important for their actions but many important sources of their shared language and patterns of explicit and implicit rules, are received through partaking in the activity. The meaning of one's actions with others are, to a large extent, received and inherited; the communal life-world is the source which allows these individual intentions and plans to be formed, not the other way around.

The phenomena of "falling into" patterns of habits, that we mentioned earlier, gives us a clue to understanding how we continue cultivating certain ways of life. Several of our actions are to some degree shared if interpreted in light of the larger horizon. Shared action is a "generally unspoken willingness to be in company with another for a certain length of time" as Moran writes (2021, 4). I may fall into the role of being a willing guide during a hiking-trip in the mountains. Being a guide is a temporary, shared social role with its own distinctive set of demands and expectations imposed by the nature of the relation to the other hikers. Similarly, in going to the store with someone, I defer to the one who has the intention to buy something. I am alongside but also a follower (similar to a dance partner).

Margaret Gilbert's discussion looks at the phenomenon of walking together and she highlights notions such as rights and correlative obligations of the participants to one another, grounded in the activity itself. However, it does not seem obvious that these concepts are at all needed to describe the phenomenon in question. Contractual language and talk of "rights" rather seem to fit into a certain philosophical vocabulary belonging to a certain received tradition.

Although she does acknowledge that there are “pertinent background understandings involved”, she does not elaborate and thus frames the problem of joint action in terms of social contract theory and “concurrence conditions” rather than a more implicit, communal and tacit belonging to a social world that enables certain normatively constrained practices (2021, 4). Gilbert talks of “background understandings” without the presence of explicit agreements, but tends to see these as conventions. I believe this way of setting up the problem is reading explicit “agreements” into the tacit communal social situation, projecting a certain juridical and political vocabulary onto the social scene.

Even taken-for-granted social forms such as “going for a walk together” are historical cultural forms (one finds accounts of such forms in Plato’s dialogues or in Rousseau). As Moran points out, there are differences between a leisurely stroll, a trip to the shop, a mountain hike, a brisk walk for exercise, a reconnoitre of a new neighbourhood, and so on. Each social form has its own implicit regimen and normative structure, and our conceptual vocabulary needs to be rich enough to do them justice.

Gilbert’s account certainly highlights some important aspects of joint actions. However, it is not at all obvious that the concepts are derived from the isolated analysis of the small-scale case itself, rather than being an imposed interpretation of the phenomena that presupposes a whole background philosophy, in this case the contract theory. This philosophy involves a host of inherited concepts that is best understood, not by looking at her isolated case (where the conceptual vocabulary is taken for granted), but in the light of the history of political philosophy, by looking at all the important thinkers and writers partaking in the conversation through historical time, concerning contract theory, commitment, and fundamental rights.

Social theories, therefore, need to acknowledge their situatedness in history, that their interpretation of the human history, itself belongs to cultural, political, and philosophical traditions and practices. The central concepts and grounding metaphors of any social theory belong to the larger enveloping life-world that is a dynamic, continually changing and essentially temporal, historical cultural form, as Husserl and Heidegger emphasize. “Being-in-the-world,” as Moran writes, “is essentially historical” (2021, 5). This means that to philosophically make sense of our joint action and cooperative activities, we cannot ignore this larger pattern of history.

We can say then that the individuality of a person is developed and realized socially through time; the self emerges from participating in a creative appropriation, interpretation and challenging of the boundaries of the inherited communal life-world. If we want to develop and build our character, we need to participate in practices and actions and find solidarity in the new

groups that are constituted by these practices. For example, we can participate in shared acts of “resistance” to the dominant culture, by engaging in writing poetry, philosophy, or political action. For the poet, even if he places the words to paper in solitude, he is still engaging in the collective tradition of language and thought, and he is unavoidably contributing to a historical dialogue of poets. Similarly, by engaging in the activity of philosophy, we become part of an intellectual community and with it, a conversation between philosophers through the ages (2021, 5-6). By engaging in these joint activities, and thereby participating in the communities that are co-constituted by them, philosophers can develop a sense of individuality, making something unique out of what is common and shared.

Thus, by looking at these more complex cases of shared activities, Moran argues that all social groups, families, groups, institutions etc., are integrated in fields nested into a unified, total, and essentially temporal and historical intersubjective world. This is what Husserl calls the “world of spirit,” which is imbued with “communal spirit” (2021, 6). Each person lives as “co-persons,” participating in a shared consciousness, living in an open horizon of others in society (2021, 6).

The communal world of spirit, or group mind, is often ignored by current Anglophone discussions of collective agency. John Searle, as I mentioned in chapter 2, simply mocks the very idea. This fact seems to indicate that the discussions concerning joint action and social ontology are actually unavoidably shaped by deeper questions concerning philosophy of mind and metaphysics. For phenomenology, we all implicitly or explicitly belong to a “communicative community”, a shared social horizon within which group actions make sense. Consequently, the intersubjective community becomes ontologically primary; it is the relational and intersubjective realm that makes the subjective self-consciousness possible and not the other way around.

Thus, by participating in shared actions and attending the same things with others, we can start, through a long process of education and maturation, to become individual agents of our own action. Also in a mature setting, with agents that are capable of rational reflection and deliberation, we remain dependent on others, on different forms of social groups, to perform our individual actions. In friendship, for example, our ability to act is enhanced by the help of our friends. Indeed, this family of phenomenological social thought is not individualistic, but the individual uniqueness of a person does not disappear by belonging to a larger communal spirit. In certain contexts, in the camaraderie of soldiers, of revolutionaries, or in a community of philosophers, the group can be tightly “connected” without necessarily being oppressive to the individual. The degree of the “fusion” or “connectedness” does not decide whether the

community is oppressing or erases individuality. This is decided, rather, by the *the moral content* enacted and expressed by the actions of the group. Some tightly knitted groups are healthy and enhancing for individual agency, others are oppressive.

Most contemporary accounts of joint action under-emphasize the degree to which the collective and anonymous worldly scheme of practical knowledge sets the parameters for action. This worldly context (the 'life-world') is most emphatically not just a set of beliefs, whether implicit or explicit. We participate in a pre-predicative world with temporal horizons of past and future, a storehouse of possibilities and affordances.

Moran agrees with Heidegger that this "worldhood-belonging" is apprehended first and foremost through mood. Moods are world-disclosing (2021, 22). They simply befall us, and often do so while we engage in activities with other people. We are always in a mood, but where we are, and what social role we play in the communities we belong to, will make certain moods easier to fall into than others. Even casual everyday normality is a mood although usually only noticed when disturbed by another "counter-mood". Moods disclose the world as significant: "This enmooded worldliness is a fundamental ingredient for joint action but it rarely featured in the current discussions. Yet without this 'canopy' joint action is impossible" (2021, 23). Thus all begins with the priority of the social and communal and recognizes a kind of "we-consciousness" that supports the individual beyond his or her I-centeredness and provides the possibility of encountering the other as another "I" (2021, 23).

We have argued that in order to develop a richer phenomenology of joint agency, embeddedness in the all-encompassing horizon of the life-world cannot be ignored for the purpose of building a social theory out of simpler social cases. One can in fact identify a hierarchy of acts, connected to social roles and ways of life, that are in turn nested in the practices and meaning-creation of communities. Most daily actions are not based on explicit rationality or even partial plans but are customary. The social world with the "alter egos" in it is arranged around the self as a centre in various degrees of intimacy and anonymity (2021, 24). One could object to this, following a thinker like Bratman, that "planning" need not be explicitly formed or even wholly conscious, but can in most cases function as nondeliberative habits. To this, I will respond that, although the concept of "planning" has its own connotations and specific limitations; applying it as the most fundamental concept that makes sense of cooperation and joint action, might not be sufficient. For the phenomenologists a richer, more dynamical, and historically minded philosophical vocabulary is needed to make sense of all the different forms of social groups, ways of joint attentions and cooperation. Here, vocabulary specific for irreducible social phenomena is essential for understanding how common action

takes place inside a large network of mostly anonymous, hidden, communal frameworks of unarticulated, practical, and habitual meanings (2021, 24).

Chapter 5: Joint action and communion

5.1 The linguistic dimension

Let us return to the question of whether our actions can be explained by referring solely to our individual intentions, desires, and beliefs. Bratman's reductive planning theory, we remember, presented quite strong arguments in favour of an individualist view of group intentions and actions, that we need not presuppose a particular irreducible social phenomenon in order to make sense of how individuals cooperate and form group intentions and perform actions. I found the we-intentions of the non-reductivists to be in some senses less convincing than Bratman. The enactivists, although interesting, seemed to be attacking strawmen with inadequate critiques of the standard mind-centred theories. Even though Bratman's theory seems to be the most developed and convincing account, his individualist approach seems inadequate when looking at a deeper level of social reality. Consequently, my answer to the introductory question is negative as I do believe there is a need for a more comprehensive account than Bratman's. Patterns, habits, institutions, history, and shared forms of life need an account of the underlying intersubjective life-world that makes human cooperation and group activity possible. This was provided by Moran and the classical phenomenological tradition. Hence, two different approaches, one individualist deriving from Bratman and one holistic represented by Moran and phenomenology, provide insights that are indispensable in order to understand social and group action. Are these theories inevitably at loggerheads, given the fundamentally different grounding assumptions, vocabularies and proceeding methods of each philosophical approach?

I will argue that there is a way to integrate insights from both. To do so, I will turn to Charles Taylor's social theory. Although his philosophy doesn't include a systematic exposition of the concept of joint action, he does give a central role to joint attention in *The Language Animal* (2016). To integrate elements of both Bratman's theory, as well as crucial insights from the phenomenologists, it is helpful to study a thinker such as Taylor who has been dedicated to make sense of both traditions.

To begin with, I will address Taylor's "expressivist" approach to meaning and language. This theory includes elements in common with the enactivists, emphasizing embodied expression as central in interaction. Unlike the enactivist, however, it recognizes the indispensable role of mental properties in human action and practice. Taylor furthermore builds on the work of Tomasello (1999) and lastly the phenomenological tradition. The theory

connects questions concerning the social to those concerning language. Language is here understood to be “all-inclusive” and holistic; the linguistic dimension, as Taylor terms it, includes all forms of human communication and expression, the whole range of meaningful media. Character, habit, body-language, gestures, facial expression, the creation of music, art, dance, all these “symbolic forms” through which human beings are able to express meaning, are included. Language is thus not understood as designative, that is first and foremost representing ideas or things in the world. Rather, language is connected to the meanings we express through action and practices, making manifest “the significance things have for us in virtue of our goals, aspirations, purposes” (Taylor 1985, 218).

To understand the meaning of “expression” in this context, let us look at the following scenario: Imagine that you are taking a walk in the neighbourhood on a pleasant day in spring. Then imagine that you see someone approaching you from the other end of the road. It turns out to be a good friend. You spot a smile on her face and greet her by waving your hand; she greets you back. It is already apparent to you from a long distance that your friend is in a particularly good mood today. This is because you know your friend very well and you know that she is bad at hiding her feelings. Something is immediately expressed through her smile: the mood of joy. This mood becomes evident from a relatively long distance because it is embodied in such a way as to be made manifest. “Manifest,” in this setting, means being directly available for all to see. The mood would not be manifest if there were only signs of its presence from which you could infer that it is there. If for instance your friend has put on a nice dress, you might *infer* that she is on her way to a party. This fact would however not immediately be manifest in the expression itself. The “I’m going to a party this evening” is not immediately manifest in the expression (the smile), but the mood joy *is* manifest in this case.

The example of facial expression is useful to grasp the concept of expression as understood by Taylor. If your friend have an expressive face, you might see joy and sorrow directly. There need be no inference as the moods are manifest, in the only way they can be manifest in public space. Contrast this with another case in which you visit your brother. He tends to hide his feelings; he has a “poker face,” but since you know him very well you also know that whenever he feels angry, he starts blinking. This day, you observe that he is blinking more than usual and from this you infer that he is angry. In this case, the eye blinking does not amount to an expression of rage because it is not directly manifest in the expression. Expressions, in Taylor’s sense, make our feelings manifest by embodying them.

Expressive theory maintains some of the mystery surrounding language, in contrast to philosophical theories that are, either unconsciously or programmatically, “anti-mystery” or

“anti-metaphor” (Taylor 2016, 105). The latter theories take the designating relation between word and objects, or in more contemporary terms, between sentences and their truth-conditions, as the logical basis of language, and hope to illuminate meaning by tracing these correlations (1985, 221). Taylor in no way denies the existence of the designative dimension of language; indeed, he understands it to be vital to scientific rigour. However, he insists that to understand how human beings live and act as individuals and as social and cultural communities, we need to consider the full shape of the human linguistic capacity.

How is this important to understand action and group activity? According to the expressive theory, when a person or a group express meaning through their actions, then this meaning cannot be fully separated from the medium (the form in which the meaning is expressed), because it is only manifest through this form. In the small-scale case of sociality, in a conversation for example, it is difficult to make sense of the meaning of the conversation without looking at the actual way those meanings are manifested through facial expressions, gestures, metaphors, images and narrative anecdotes. Take away these media of expressions, and one quickly loses the meaning of the conversation tout court. In consequence, human actions give expression to meanings that cannot be explained by these meanings being related to something else, to an isolated object or a representation, but must be understood in the context of a wider horizon of shared communal sense-making. If the meanings expressed always participate in a holistic horizon of significance, it follows that the method of isolating terms and tracing correlations cannot work for expressive meaning (1985, 221).

Another important feature of this theory is that the paradigm-expressive objects function as wholes. If we look at a face, or read a poem, or contemplate a work of art, we cannot simply break the objects down into parts, or show the whole to be a simple function of the parts, if we want to understand what is expressed (1985, 221). This also goes for the expression of “social wholes”; in order to apprehend the meanings that are manifest in an opera performed by a group, we need to interpret the expressive actions and singing words as a narrative whole.

Thus, an expressive account of meaning cannot avoid subject-related properties. Expression is the power of a consciousness (or several); and expressions *manifest* significance and hence essentially point us to communities for whom these things can be manifest. For Taylor, the horizon of meaning and the shared linguistic dimension cannot be entirely separate from one another. *The Language Animal* argues that language can only be understood if we give an account of its constitutive role in human life; this constitutive force of language is named the linguistic dimension, where the use of either words or symbols, or expressive actions,

is guided by a sense of rightness, which cannot be made simply a function of success in non-linguistic tasks (2016, 288).

If we follow Taylor in understanding language in a broad sense, including the whole range of symbolic forms (words, gestures, symbols, works of art and so on), it becomes clear that there is more than one level of linguistic communication involved in human thought and action. Taylor posits three main levels: There is (1) the descriptive level of communication which includes standard attributing properties as well as self-articulative properties, which clarifies and transforms the space of human meanings. Furthermore, there is (2) that of enacted communication, where meaning becomes manifest in embodied action and expression. Finally, there is (3) communication through portrayals and metaphorical imagination (2016, 261-62). (We can henceforth call these three levels description, enaction and portrayal.) Note that all three levels, for Taylor, have an expressive dimension; even the descriptive level, especially that which concerns self-articulation, contains an expressive dimension.

Taylor's scheme is meant to challenge the attempts of post-Fregean philosophy to validate a "modest", "mystery-free" understanding of language, one in which learning a language would be equivalent to learning to generate extensional truth conditions for its various depictive combinations (2016, 262). Again, expressive theory does not deny the validity of the designative function of propositional and conceptual language; after all, it is part of the description level. However, in the expressive theory the designative function do not have primacy over the expressive function in the social and cultural realm. For Taylor, philosophy needs the richer vocabulary of expressive theory in order to make sense of the thick cultural meanings which are vital in order to understand terms designating social relations, hierarchies, and strong evaluations regarding higher activities and ways of life (2016, 262). In the search of understanding human intention and action, we cannot bracket questions concerning the good life, what he calls matters of "strong evaluation".

Strong evaluations can roughly be described as the fundamental visions of the good that we live by. These visions involve "discriminations of right or wrong, better or worse, higher or lower, which are not rendered valid by our own desires, inclinations, or choices, but rather stand independent of these and offer standards by which they can be judged" (1989, 4). They cannot be reduced to attitudes, to beliefs, desires, or intentions, but all these are shaped by this underlying orientation towards the good. A philosophical vocabulary is needed to be able to interpret how the actions of persons and communities participate in these larger patterns formed by the strong evaluations.

One could object that such larger notions of strong evaluations cannot be manifest in the expression of actions, because they are not immediately apparent and therefore not manifest. There might be *signs* in the action that point in the direction of the strong evaluations motivating the actions of a person, even in the seemingly banal actions of everyday life (making a meal, drinking a glass of water, reading a book). But the strong evaluations about the good cannot be directly manifest in actions, and therefore must be left out from the study of action.

To this objection, it is useful to reintroduce Bratman's concept of intentions as elements in larger plans. For Bratman, we are fundamentally planning creatures, given our limits and our partial rationality (Bratman 1999, 2-3). The plans that we form on the basis of rational deliberation shape our everyday habits, social roles, and shared cooperative activities. These in turn form the basis on which our intentions, at any point, can give rise to actions. Planning is indeed essential, but on the ontological level, we are first and foremost language animals for which things have significance. We can grant that intentions often function as elements in larger plans which coordinate our habits and actions. But plans, in turn, are elements in a fundamental vision of strong evaluations, oriented towards the good. This fundamental orientation shapes our habits and actions on a more fundamental level. Thus, an individual action might not give direct expression to this overall orientation, but our actions, insofar as they are rational and can be made sense of at all, are part of larger patterns of habits, which in turn are part of a way of life. These habits and actions together might indeed give expression to the strong evaluation that motivates this life. And insofar as any individual intentional action cannot be adequately made sense of without reading it into this larger background of habit-patterns, the expressive dimension must be taken into philosophical consideration. How then does a way of life become manifest in an expression? This can be done through enacting a certain manner of speaking and acting, a certain bodily life-style to do with clothes, expressions, and so on. But it can also be manifest through literary portrayals: a certain strong evaluation becomes manifest in a novel or a poem (Taylor 2016, 319).

Understanding the role of art, poetry and novels might illuminate the expressive side of joint action and shared activities, but this is beyond the scope of this thesis. What is of importance here is how the expressive account might help us integrate Bratman's analytical reductive accounts of cooperation and joint activity to a more phenomenological holistic approach. Here we can see planning and cooperative action as forming parts of a larger set of patterns motivated by matters of strong evaluation. The expressive side of activities and action connects the latter with the overall background horizon of intersubjectivity which Moran's phenomenological account sketched for us.

The expressive scheme of the three dimensions of language, the descriptive, the enactive, and the portraying, is not only outside the purview of the “modest”, mystery-free philosophy of language, but this latter descriptive dimension could never be useful without the background activity of the creation and definition of expressive meanings (2016, 262).

What needs to be recognized, according to Taylor, is the primacy of communication, that the way human beings exist in the world is essentially dialogical. In *The Language Animal*, building on thoughts of the romantic thinkers Johann Georg Hamann (1730-1788), Johann Gottfried von Herder, (1744-1830) and Wilhelm von Humboldt (1767-1835), he argues that the primary locus of language is conversation. This means that language never just develops inside an individual’s mind, to be subsequently used in communication with others; it evolves, rather, in the interspace of joint attention, or what he terms *communion* (2016, 50).

In previous chapters we have discussed how individual and group intentions and actions seem to participate in some larger patterns of activity (see chapter 3 and 4) although we were not able to describe these any further at that stage. Using the expressive theory, however, we are closer to understanding the patterns and propose to call them the linguistic dimension, that is, the whole of language. According to the expressive conception, social groups need expressions in order to make their ideas manifest to themselves and to others. And manifesting meanings by participating in language is a form of communal joint activity that all language-speakers and expressive communicators participate in (1985, 229).

However, if everything social participates in some sort of omnipresent linguistic dimension, and if this dimension essentially is a shared activity, then has not joint action become the basis of the whole of social reality? And doesn’t this stretch the concept of joint action to the point of vacuousness? I will argue that if we understand joint actions as including shared activities and communal practices, then indeed, language may be considered a form of joint action. And this does make joint action quite fundamental. Furthermore, the holistic approach to joint action implies that particular cases of joint action and cooperative activities are able to express meaning and rationality just because they participate in a larger pattern of communal activity. Our ideas, intentions, beliefs, and attitudes thus do not properly make sense until their expressive meaning can be articulated and interpreted in the joint activity of a speech community. This also makes the media of shared expressive action a basis not only of our group intentions and actions, but also of our individual intentions and actions. The active part of language (when we try to communicate in some way), can only be deployed against the background pattern which we can never fully dominate. This doesn’t amount to determinism,

since the background linguistic dimension can never fully dominate us either, given our ability to constantly reshape parts of it, thus in some ways transforming the whole (1985, 232).

Human beings are constantly shaping language; we are able to mint new terms, displace old ones, and cultivate new interpretations of certain words. By collectively changing our fundamental expressions, symbols, metaphors, narratives, and ways of life, we are thus able to create grounds for new concepts and ideas; language is therefore constitutive of thought. The background patterns are manifest in speech, but conscious speech is nevertheless like the tip of an iceberg. Imagine two members of a band, a guitarist and a drummer, sitting in a bar, trying to write a song together. They try out different sentences, looking for appropriate rhythmic combinations, trying to find the “right” words that fit the music. Here the sense of rightness is evident. They participate in a joint activity of conversation and introspection, looking for expressions which can communicate meaning. Note that in this context the guitarist cannot have total control over his own vocabulary. He comes up with a word and mentions it to the drummer, who in turn evaluates it, and so on back and forth. Most of the words arrive quite spontaneously. Here is a shared cooperative activity taking place as both have plans to write a song together and their subplans mesh in such a way that they are able to write it in concord. They also have a shared goal and (perhaps) also a more or less articulated meaning, a theme, that they want to make manifest in the song. This joint action is consequently also an expressive action, and to understand the full meaning of it we must interpret their activity in view of the already shared background pattern of language available to a certain community and culture. Furthermore, much of what is going on in shaping the joint activity of the song-writers is not in their purview; their deployment of language rests on preconscious patterns that are sometimes easier for others to understand.

Our new coinages, phrases and even ideas, are never quite autonomous, quite uncontrolled by the rest of language; they can only be understood because they already have a place within the “web”. Language, seen as an activity and also as an intersubjective background pattern, is thus what is constantly created and recreated in speech, and the primary locus of language is the conversation. This intersubjective conversation is fashioned and grows within the life of the speech community which in turn is partly constituted by this web which we can never fully dominate and oversee. This means that language cannot wholly be the property of any one individual; it is always largely *our* language. Taylor writes:

Language is not an assemblage of separable instruments, which lie as it were transparently to hand, and which can be used to marshal ideas, this use being something we can fully control and

oversee. Rather it is something in the nature of a web, and to complicate the image, is present as a whole in any one of its parts. To speak is to touch a bit of the web, and this is to make the whole resonate (1985, 231).

Given this expressive theory, the words we use only have sense through their place in the whole web. Therefore, we can never have a clear oversight of the implications of what we say at any moment. This is also the case for our actions, insofar as these express meaning and strong evaluation. When the two band-players write a song together, they might have a shared plan, they might satisfy all the conditions sketched by Bratman, but if we are to grasp the meaning of their shared action/activity, we need to consider what they try to express in the action, what they try to make manifest in words put to music. Our language, even for two highly intentional cooperating partners, is always more than can be encompassed in words (1985, 231). But how exactly does the expressive side of language become embodied in individual actions? To understand this, we need to turn to Taylor's account of joint attention and how this might be related to meaningful action.

5.2 Joint attention and communion

If language is essential in the background patterns that influence our ability to participate in shared actions and practices, what does that tell us about becoming part of social groups acting together? Here we must explore another aspect of the issue of joint action (which we also discussed in chapter 3), namely joint attention. Regarding ontogenesis of language, when and how the language develops in the organism, it seems language can only be imparted from within relations of shared emotional bonding and joint attention, what Taylor terms "communion" (2016, 55). This means that in order to partake in language, and to act meaningfully at all, one must first become part of a community of family or caregivers. Within the expressive theory, human beings are not individuals first who later on come together to perform an activity and become part of a group. On the contrary, it is only because an individual is part of a speech community that it can later on develop a sense of individuality. The communal "we" dimension is thus more fundamental than the "I" dimension. In contrast to some individualist models, Taylor denies that language can be generated from within. It can only come to the child from her interactive world, her parents, and her surrounding milieu.

How then are children drawn into the linguistic dimension through the caregivers? The relationship between joint attention and joint action is important to further understand this. If language is a fundamentally shared activity, then the way into participating in this activity, is

through bond-creating joint attention or communion. Taylor is here inspired by Tomasello who views the ability to grasp the communicative intentions of others through partaking in joint attention as the crucial new capacity which allows human children, in contrast to other animals, even advanced primates, to become language users (2016, 55).

What does communion mean? It means to be joined together in the sharing attention towards the same object. In the case of joint attention as communion one is attentive to the same reality in a state of intense emotional bonding (2016, 62). Thus, for Taylor in this context, the concept of communion encompass an intense form of joint attention, where those involved become deeply “porous” to the influence of each other. Contrast this with a case where such porosity is usually relatively absent: Imagine some co-workers coming together in the lunchbreak to watch a football match. Here everyone pays attention to the same game, and it is indeed a way of being together and of bonding. But the colleagues (given normal circumstances) would not risk becoming intensely porous and vulnerable while watching the match; they are certainly bonding, but at the same time they want to protect their autonomy and individuality from becoming too emotionally and existentially influenced by the others. In a case of communion, especially between children and parents/caregivers, the joint attention is more intense, the “porosity” allows a deepened bonding and perhaps even “fusion”. Communion is the formation of fundamental bonds, where the sharing in consciousness, emotion and attention happens simultaneously.

But a baby and a mother, one might retort, surely cannot share consciousness or be of “one mind” in a literal sense? It is true that the child cannot access her mother’s habits, concepts, memories, and skills directly through communion, but it nevertheless seems plausible that the child experiences little distance between itself and the mother, and that the sharing of attention, emotions and even purposes, is exactly what makes it able to enter into the linguistic dimension, which we can interpret, following classical phenomenology, as our shared lifeworld. Research also suggests that mothers, in most cases also feel an unusually intense connectedness to their babies, as if they were still a part of themselves. Thus, in the case of successful communion, we may speak of sharing one “mind” or “consciousness.”

If we thus assume that consciousness is something that can be participated in- as well as something that can *in-fluence* (“in-flow”) into the consciousness of others, then the intense relationships that are created through joint attention and communion in the early years of childhood, is a plausible way for the human child to enter the shared intersubjective linguistic dimension. This “sharing-nature” of both consciousness and language seems to be a basic feature of human life from the beginning and throughout life.

Here, expressive theory leans on Tomasello who argues in *The Cultural Origins of Human Cognition* (1999) that joint attention is essential for the acquiring of culture and language. For Tomasello, chimps and other higher mammals can identify their conspecific's ordinary intentions, can observe that one is seeking food while the other is preparing to attack. But only humans can perceive that another human wants to communicate interconnected skeins of meanings (Taylor 2016, 55). The very first formats which precede and prepare ground for imparting language create an intense common focus of attention. This common focus makes sense of the patterns of reality that are made manifest to consciousness. This means that acquisition of our first language is rooted in common attention or joint attention (2016, 56-57). Taylor thus agrees with the Wittgensteinian view of Radenovic et al. in seeing "joint attentional frames", or "referential triangles", where two speakers share the same reference, as lying at the very core of language-acquisition. This ability is understood as a human primitive and the foundation for shared action.

The child, from the very beginning, struggles towards participating with the caregiver in the activity of joint attention. Joint attention itself can be meaningfully described as a form of action since focusing on something together is something we *do*. The child learns to focus her attention on a shared object with the mother (or caregiver). She tries, for example, to reproduce the word that is addressed to her. In such a case, in early childhood, the child strives to intimate "our" word, or "the" word. The word itself becomes the object of joint attention. In the early stages of language acquisition, the focus is on "getting the word right," being able to use words when engaging in the activity of communion. The "getting it right" aspect shows that language from the beginning has normative forms, lexical grammatical. But if this is so, it seems that the shared activity of language-acquisition acquired through communion, already presupposes a larger background pattern of the intersubjective life-world. Thus, to the degree that the joint attention is successfully shared, this implies that the members of the group (here mother and child) are able to partake in this larger expressive dimension of meaning.

According to Taylor, the child's struggle to be able to use her first words is a kind of struggle for "ritual" rightness. Ritual, in this context, means that the use of the words has a sort of symbolic and expressive function. The way the words are uttered as "declaratives" ("Anne dance!", "Cat gone!") tells us that the child is using them as "vehicles of sharing" (2016, 62). The words aren't just used to describe the surrounding objects; they are used as a way of becoming part of a linguistic community. If for instance a child sees that the family cat has left the room, she may announce "cat gone!". What is most important here is that she by uttering "cat gone" is initiating, or prolonging and intensifying communion with the adult. Her

consciousness becomes part of others in the room by sharing in the same focus of attention. Furthermore, the child is imitating the adult who has been initiating such shared attention, among other gestures and actions, using words like “see the kitty?” For the child, these words enable the creation of joint attention by other means, and it prolongs and intensifies the emotional bond-making that is communion (2016, 62).

The “ritual rightness,” as we have termed it, involves the learning of certain habits while engaging in the interconnected linguistic dimension. If it is true, as the expressive theory suggests, that the parts of language (the individual words) cannot be made sense of unless connected to the web of language as a whole, then the purpose of “ritual rightness” is to be able to discern the same connections between words as the parent, connecting her small “web” to the larger “web” of the community to which the parents are parts. By exploring how “Cat gone” is the right word, not just as description but as a “trigger” enabling the establishment of communion, she shows that she is connected to the skein of meanings that connect certain words to the object of attention (that the cat has left the room).

In the same way “Anne dance” is used to commence joint attention and emotional bonding around the fact that the child can dance and is now dancing in front of the parents. At this point in the process the child is already out of the animal zone and has entered the linguistic dimension. Even apes raised in human environments cannot be initiated in the same way as described above. Tomasello comments on the difference: “basically all their [the apes’] production ‘of signs’ are imperatives, to request things, to the neglect of the declarative sharing of information” (2005, 290). The apes thus can use signs to get things they desire, but they cannot use words as a way of participating in communion. Consequently, the realm of the linguistic dimension is closed to them. Human beings at an early stage can engage in a practice, which Taylor terms “protointerpretation” (2016, 63), originally carried out in close communion with the parent/caregiver. This practice helps the child in defining and redefining her desires and longings in order to live with the pattern of fulfilments and frustrations she undergoes throughout life. This later on develops, into a full-blown “interpretation of the condition of man in the world,” an unending task of “finding the meanings which can make sense—bearable sense—of our lives” (2016, 63).

Thus, a human being, according to expressive theory, enter into the linguistic realm through shared attention and communion. From there they continue throughout life in the shared activity of interpreting existence. This is largely achieved through the practice of conversation and by engaging in joint action and cooperative activities, which through time, helps create and further develop social groups and institutions.

5.3 Expressive theory and joint action

In light of the above, we can now try to connect the expressive theory with insights from the theories of joint action discussed formerly. For expressive theory, group activities must be made sense of in light of the full range of expressive modes, which Taylor after Cassirer calls the “symbolic forms.”¹⁰ From this perspective, we can criticise the standard “Cartesian-derived monological approach” that puts self-awareness and isolated hidden “mental states” prior to our eventual access to an intersubjective world. Instead, we propose a “plurality of the I”; where the individual person fundamentally contains a community as a whole, in order for it to be a person at all. Abstract away the individual self-consciousness from his friends, his family, his place, his groups, communities, and history, in fact the shared linguistic dimension he is part of, and you end up with an empty shell. This empty shell might be useful to discuss as a concept at times, but to make sense of the rich meanings expressed in our actions, particularly the life and activities of social groups, understanding how individual thoughts and actions partake in the larger “skein of meanings” of the communal life-world is vital.

The interaction with the rest of reality is contained and reflected in the individual consciousness. Groups, communities, institutional practices, and the languages that are co-constitutive within them, together form a particular and concentrated forcefield that gives life to this self-consciousness and its personality. This seems to make expressive theory both holistic and collectivist, in that it makes community and intersubjectivity ontologically prior and more basic than the individual person.

A relevant argument directed against expressive theory could be stated as follows: Taylor’s collectivist social ontology cannot make sense of the individual mind’s “hidden depths” and “interiority”. He reduces the role of autonomous rationality and intention in cooperative activities by mystifying the social linguistic realm, casting it as some unexplainable mysterious whole. Furthermore, if individual expressions are ontologically so tightly connected, how can we make sense of lying, illusions, and the fact that individuals hide our intentions and desires all the time? Too much emphasis on the irreducible social and shared, underestimates the vast amount of deceit, lies, and concealed conflicting wills in social reality. In the case of groups, are there not many cases in which the individual members have quite different opinions and intentions than the official opinions and intentions of the group as a whole? Clearly, our ability to disclose our inner thought or lie and hide our intentions and

¹⁰ Ernst Cassirer, *The Philosophy of Symbolic Forms* (1953).

meanings is one of the defining features of human social life. Consequently, we should start by understanding the individual action in modest social settings before moving on too quickly and lazily to a “collectivist reductivism” in which human individuality is nothing but the expression of some historical communal realm of language. We as individuals might have the capacity to form “we-intentions” and a sense of shared life, but the fact that these intentions always occurs only in individual minds should not be obscured by a communalist notion of the plurality of the I or the “we” of the self.

I think the above objections misunderstand Taylor’s “collectivism” and present us with a false account of non-individualist theory. Embracing the ontological priority of the intersubjective life world, the concept of the shared space of conversation and language and emphasising that we always already partake in a shared historical consciousness, none of this mean that there is no place for the notion of idiolect. On the contrary, our grasp of language is complex, as complex as conversational exchange, which requires that each person develops its own way of manifesting the expressive meanings through actions. We use and participate in “the” language, in a collective linguistic dimension, but we do so in different ways; we come at it from different directions. This sense of the difference of perspective is essential to normal conversation. The intersubjectivity of language doesn’t presuppose uniformity of mind; it presupposes that a conversation or conflict can only take place in a shared “life-world”. If there was no shared underlying “whole” of language, coming together in a conversation or in a conflict, as a group, would be impossible. We try to put in a way which our interlocutor can grasp, but this presupposes that we to already, to a large degree, share a collective life, more than we might be aware of.

Children grasps this intuitively very early. Their address differ between different interlocutors as they test the adult’s response to them (2016, 59-60). In joint attention, there is thus an element of play that is crucial for the children to experiment and to develop its own “take” or “style” in relation to the shared linguistic realm. It is in fact because Taylor understands the subject as *in itself* an undisclosed depth, that he can maintain the fact that the individual partakes in the collective historical consciousness of a community (1985, 94). However, language as well as community cannot be totally transparent. It is the inexhaustible and indispensable interiority of the individual self-consciousness that is always also given in group activities of expressive communication, made possible by life in common.

The original interiority of the mental cannot be dissolved to the exterior world of groups, this is true. But the reason for this is not individualism, the focusing on the lone individual’s intentions, desires, beliefs, and attitudes. Rather individual mental depth is possible because

groups and communities also contain a depth and an inexhaustibility. Thus, the individual expressive mind is hidden and manifest at once because the community he or she partakes in is hidden and manifest at once. To further understand this, we will look at the reality of language and how it manifests during a conversation in a group. A person does not have at her disposal, in the mind, a representation of the whole of her present vocabulary of concepts, metaphors and images. And yet, when engaging in a conversation, words manifest themselves in a sequence, or many sequences, such that it becomes clear that she indeed possesses a vocabulary larger than what is immediately transparent to her self-consciousness. The whole of language, the background pattern, is never completely manifest; like the consecutive notes in a musical sequence, some words must vanish as the next ones appear, for the utterance to make sense as a whole. To communicate every possible combination of sentences at once would be impossible and it would not be communication. And so here we see that language, even though it is a complete “web,” it is partly manifest and partly hidden at once.

In a similar manner, the “we” that is implicit in the “I” is not a social concord that simply dispels the hiddenness of the individual. Getting to know a person requires the person to disclose some of her relationships to different communities (family, friends, workplace, political community etc.) but not all at once. Trying to list every single intention, attitude and desire that constitute her personality would not work either, since, as with the sequence of individual notes in a piece of music, some things must be disclosed and then vanish in order for the conversation to make any sense and for the meaning to become manifest as a whole. And if we understand conversation as a fundamental form of joint activity it becomes clear, I believe, that the expressive dimension of language must be accounted for in order to access whether the conversation is meaningful and rational. A completely meaningless conversation would be one in which no meanings were manifest to any of the participants, that is, it would be impossible to interpret the individual words in connection to the larger linguistic dimension. Even a conversation that is done for the sake of pure play, with no external purpose other than having the conversation itself, would not, strictly speaking, be meaningless in the mentioned sense, because the words as a whole manifest a form of meaning, the meaning of joy in pure play, improvisation, and the sharing of attention towards certain words. Such a conversation would still be a meaningful “inter-action” taking place in shared space of an intersubjective communal life-world.

Let us go back to looking at how all this might be connected to the insights from the former chapters in this thesis. From non-reductive theorists, such as Searle, Gilbert, and Tuomela, we learned that there seems to be some inherent social capacity or an irreducible structure that

gives us the sense of the “we” in our group activities. For them, joint action need an irreducible social “dimension” that is participated in among the members of the group, for the action to be really intentionally shared. However, we did not see the need for the biological “we-intentions” of Searle, nor any strict need of the concept of plural subject in order to make sense of joint action. We thus agreed more with Bratman and the reductivists in that joint action and shared cooperative activity need not presuppose any “special” kind of capacity to form we-intentions. However, I disagreed with Bratman as I argued that joint action seemed nevertheless to participate in a deeper irreducible social dimension. Based on the previous discussion, I argue that this social dimension is what the phenomenological tradition called the intersubjective “life-world.”

Furthermore, what we learned from the enactivists, the emphasis on embodiment and on interaction in spontaneous group activity and joint attention, was that joint action might need a richer conceptual vocabulary in order to make sense of the shared activity, particularly the activity of conversation. But the enactivists came too close to behaviourism in trying to deny the importance of intentions and of the role of the inner life of the mind, when understanding group action. It seemed that the perspective/s introduced by Moran and the phenomenological tradition was better able to make sense of how consciousness participate in the irreducible social dimension of the communal life-world. But how can we connect the “life-world” to the reductive and individualist theory of joint action, provided by Bratman?

Taylor’s expressive theory, I believe, can provide a link between action and the life-world using the concept of the linguistic dimension. For Taylor, language includes all forms of communicated meaning, through our actions, habits, character, and ways of life. Following Taylor, we can argue that Bratman’s scheme, although ingenious in its detailed mapping of how human beings are able to act together, can be integrated into a richer social theory which knits language and community tightly together. It is because language and the community of language, that is, language as a whole and the shared “life-world,” are mutually constitutive of each other, that we can make sense of the rich plurality of different groups and institutions. These have different strong evaluations and meanings that make sense of their activities, actions and ways of life, as well as how they are able to mesh their goals, plans and ideals in concord.

5.4 Further considerations

We have here tried to connect the expressive theory to the field of joint action, by making use of Taylor’s rich vocabulary of the social world. The expressive theory, as we have interpreted

and further developed it, inverts the post-Cartesian and individualist priority; self-awareness and individual intentions should be seen as emerging out of a prior intersubjective sphere, and this is precisely what the notion of communion implies (2016, 65). All children (given the circumstances of growing up in a functioning family and community) are first inducted, not into the parent's view or her own view, but what she experiences as "the" view of the world. This view is imparted along with the language and developed within the ambit of an emotion-infused joint attention or communion (2016, 65). The development of the self comes after the constitution of the common world of joint attention, and it is only within this common world that we are gradually able to develop a fuller sense of self by sorting out different perspectives, different takes on the language and reality.

One could criticise this theory at least on two points. Firstly, if the expressive meaningful dimension of action is so important, and if embodied manifestation is primary, how do we explain and understand communication through social media and the internet? We live in an age of "hyperconnectivity"¹¹ a time in which the internet and social media connect people from different groups, with different cultural, ethical, and political strong evaluations, in a shared globalised world (Brubaker 2023, 22). Does this not demonstrate instances of joint action where the "thicker" cultural expressions that are made manifest in embodied and local group activities, need not be considered? Let us say that the students A, B, and C living in different continents of the world meet on social media and plan to meet on Zoom to cook a meal. All three are cooking their own meal in their own respective apartment, but they are still performing a shared cooperative activity that meets the criteria of Bratman's scheme: their subplans mesh, their intentions are in concord, they are all intent on doing their part, and helping to coordinate each other's behaviour, and they all share these facts as common knowledge. Do any shared horizon of meaning need to become manifest in their shared action? It seems not, it seems that this richer account can after all be put to one side. They might have their own subjective interpretations of the meanings of their action, but strictly speaking, these meanings need not be manifested and shared by all three for there to be far-off-distance group activity.

To this I respond that the hyperconnectivity of the internet-age, rather than weakening the plausibility of the expressivist theory, on the contrary makes it more plausible. Let us look again at student A, B, and C. In order for them to be able to communicate at all at such a distance, they need to participate in some form of shared language community. Perhaps they

¹¹ For an illuminating discussion concerning the social and political implications of hyperconnectivity see Roger Brubaker's *Hyperconnectivity and Its Discontents* (2023).

are able to communicate because they all speak English. It seems that a connection through the linguistic dimension is particularly important in the context of social media. Considering the manifestation of meanings in their actions, it seems that in the case of cooking together, they are after all able to communicate a sense of bodily presence through the camera. But, as many who have participated in online conversations, shared video-game activities¹², etc, the risk of deeply misunderstanding the meaning of the group members' deeds and actions, as well as the cultural ideals that motivate them, is high and often leads to discord and conflict. In a globalist, pluralist world, it is even more necessary to cultivate the ability to interpret the actions of groups and communities in light of the strong motivations and "thick" cultural and historical background of the actions. As Taylor writes: "[Even] with fluidity and change—perhaps especially with this fluidity and change—being a speaker of a widespread modern language requires sensitivity to the meanings and footings that underlie these registers [of the different social groups] and their shifting boundaries" (2016, 331).

This is why Taylor inverts the traditional priority of self over intersubjectivity. The latter, what we have called communion, comes first. When alienation from a group or community occurs and one's original sense of common life is broken, this can, although unfortunately not in every case, motivate attempts to renew and recover the social bond, and the meanings which are central to it. These are reconciling joint activities that Taylor calls "restorative conversations," which urge us to realize, again, how deep and fundamental conversation is for social groups. We *are* a conversation, and this not only points to the beginning of speech in communion, but also to the "we" of communion. This renewal of communion where people are recurrently recovered out of alienation and conflict by restorative exchanges, is especially important in our age of hyperconnectivity where alienation and loneliness may otherwise lead to heightening of conflicts.

The second point of critique I will direct at Taylor's theory, is that it does not develop an adequate ontology and philosophy of mind that can make sense of phenomena of communion and the linguistic dimension. At the end of *The Language Animal*, he seems to hint that this conception needs an updated version of the idealist, Romantic and Christian Platonist conception of an Infinite Consciousness underlying and supporting the shared space of meaningful social patterns (2016, 343-44). But he never tries to defend such an idealism, and given the contemporary context of philosophy of mind, where functionalist theories are

¹² For a discussion on the implications of Avatar-mediated spaces for understanding joint action see Mariano Pugliese & Cordula Vesper (October 2022).

dominating, it would be difficult to do so¹³. But without such a grounding ontology, is not Taylor's expressive theory vulnerable to anti-dualist critiques? Can the notion of the linguistic dimension be made sense of purely on a phenomenological level? This problem needs further discussion and development and may involve connecting the expressive theory to discussions of philosophy of mind, although we cannot provide it here.

¹³ For a discussion of the plausibility of the functionalist position in contemporary philosophy of mind see Jaegwon Kim's "Mental Causation" in *The Oxford Handbook of Philosophy of Mind* (2009, 46).

Conclusion

From the standard non-reductive theories, we have seen that there seems to be something irreducible about the social structures that allow us to act in concord. However, they base this on a particulate capacity to form a “we-intention” which enable us to act in “joint commitment” as a plural subject. Although joint action seems to presuppose a strong sense of “we”, I see no need not construct a special “we-mode” or “we-intention”. Consequently, I find myself more sympathetic with Bratman’s planning theory, stressing the continuity between individual future-directed intentions and the collective interrelated plans that make possible joint action and shared cooperative activity.

Although I agree with Bratman who claims that joint actions are explained by our individual intentions, desires, and beliefs, I found that something more fundamental than the “we-intention” or “joint commitment” was required to make fully sense of joint action. I have furthermore studied the enactive theory of Gallagher, Fantasia, De Jaegher, Fasulo, and Fuchs which connects joint action to the embodiment and enacted meanings emerging through interactions between group members. They described joint attention among children and parents, and found that even children with reduced capacity to form intentions and concepts, can cooperate in joint activity through embodied interaction.

In the further study of this interaction and the patterns that make it meaningful I sided with Moran and the phenomenology as well as Taylor in arguing for an irreducible background horizon of the life-world and the linguistic dimension. By integrating these latter approaches with those of Bratman, joint action could be shown to be connected to skeins of meanings that made sense of our strong evaluations. Similar to the interactive theory of the enactivists, we have seen that expressive theory also views human consciousness as derived from a process of embodied interaction with the environment and through participating in practices. In Taylor’s theory, however, language itself is the most fundamental level encompassing all forms of communication from the enactive to the conceptual and intellectual.

Furthermore, I argued that the matrix of language is conversation, and this remains so throughout human history, being fundamental to all activity and group intention. The concept of communion helps us understand that language comes to us through a fundamental bonding and joint attention. This communion initiates us into the historical conversation where language is maintained, altered, and renewed (2016, 58). What a particular person comes to understand

as her own position, her own existential perspective in the conversation, is a precipitate of the original joint attention or communion between caregiver and infant (2016, 65).

From this, I ultimately find that we need to complement the standard accounts of joint action with insights from classical phenomenology and Taylor's expressive theory. The latter illuminates the importance of metabiological meanings of individuals and groups, *the strong evaluations*, in order to grasp what holds them together and make them able to act. These meanings concern goals, purposes, and discriminations of better or worse, and they cannot be defined in terms of objectively recognized states, but must nevertheless be understood for a philosophical account of the social group to be complete. To understand a community or group, or a member of a community, the theoretical analysis must be complemented by an interpretive dimension that can account for meanings expressed in group activities and practices. To address this dimension, we need to understand which are the shared goods of the community, what is its understanding of the good and what kind of narrative does it embody? The traditions, the strong evaluations, and beliefs of a family, say, can be radically different from those realized in a political community or a religious community. In order to understand a particular group, we need to get inside the group's language of self-description. We should aim at achieving a truthful interpretation of the meanings of actions performed collectively by human community in history. This kind of interpretation remains an ongoing theoretical task: the joint action of philosophical conversation. We can thus conclude by citing, as Taylor himself does, the famous line from Friedrich Hölderlin: "Since a conversation we are / and hear from one another" [*Seit ein Gespräch wir sind / und hören voneinander*] (2016, 58).

References

- Alonso, Facundo M. 2020. "Reductive Views of Shared Intention." In *The Routledge Handbook of Collective Intentionality*, edited by Kirk Ludwig Marija Jankovic. New York, US: Routledge.
- Bratman, Michael E. 1992. "Shared Cooperative Activity." *The Philosophical Review* Vol. 101, No. 2.
- Bratman, Michael E. 1993. "Shared intention." *Ethics* 104.
- Bratman, Michael E. 1999. *Intention, Plans, and Practical Reason David Hume Series: Philosophy and Cognitive Science Reissue*. The United States: CSLI.
- Bratman, Michael E. 2014. *Shared Agency: A Planning Theory of Acting Together*. New York, US: Oxford University Press.
- Brubaker, Rogers. 2023. *Hyperconnectivity and Its Discontents*. Cambridge, UK: Polity Press.
- Cassirer, Ernst. 1953. *The Philosophy of Symbolic Forms*. New Haven, CT, US: Yale University Press.
- Chant, Sara Rachel. 2020. "Collective Action and Agency." In *The Routledge Handbook of Collective Intentionality*, edited by Kirk Ludwig Marija Jankovic. London and New York: Routledge.
- Gallagher, Shaun. 2010. "Joint attention, joint action, and participatory sense making." *Alter. Revue de phénoménologie* 18.
- Gilbert, Margaret. 2006. *A theory of political obligation: Membership, commitment, and the bonds of society*. UK: Oxford University Press.
- Gilbert, Margaret. 2020. "Joint Commitment." In *The Routledge Handbook of Collective Intentionality*, edited by Kirk Ludwig Marija Jankovic. London and New York: Routledge.
- Jaegher, Thomas Fuchs Hanne De. 2009. "Enactive intersubjectivity: Participatory sense-making and mutual incorporation." *Phenomenology and the Cognitive Sciences, Springer Science + Business Media B.V.* 8.
- Kim, Jaegwon. 2009. "Mental Causation." In *The Oxford Handbook of Philosophy of Mind*, edited by Ansgar Beckermann Brian P. McLaughlin, Sven Walter. Oxford, United Kingdom: Oxford University Press.
- Ljiljana Radenovic, James Connelly. 2009. "How Wittgenstein can help us understand autism better: Concepts, rule following, and the nature of explanation in developmental psychology." *Res per nomen II (langue, référence et anthropologie)*, Reims, France

- Mariano Pugliese, Cordula Vesper. October 2022. "Digital joint action: Avatar-mediated social interaction in digital spaces." *Acta Psychologica* 230.
- Miller, Raimo Tuomela and Kaarlo. 1988. "We-Intentions." *Philosophical Studies* Vol. 53, No. 3.
- Moran, Dermot. 2021. "The phenomenology of joint agency: the implicit structures of the shared life-world." *Phenomenology and the Cognitive Sciences*, Springer
- Searle, John. 1990. "Collective Intentions and Actions." In *Intentions in Communication*, edited by et al. Philip R. Cohen. US: MIT Press.
- Taylor, Charles. 1985. *Human Agency and Language: Philosophical Papers I* Cambridge, United Kingdom: Cambridge University Press.
- Taylor, Charles. 1989. *Sources of the Self: The Making of the Modern Identity*. Cambridge, Massachusetts: Harvard University Press.
- Taylor, Charles. 2004. *Modern Social Imaginaries*. Durham and London, US and UK: Duke University Press.
- Taylor, Charles. 2007. *A Secular Age*. London, England: The Belknap Press of Harvard University Press.
- Taylor, Charles. 2016. *The Language Animal: The Full Shape of the Human Linguistic Capacity*. Cambridge, Massachusetts, and London.
- Taylor, Charles. 2021. *The Explanation of Behaviour*. London and New York Routledge. 1964.
- Tomasello, Michael. 1999. *The Cultural Origins of Human Cognition*. Cambridge, Massachusetts, London, England: Harvard University Press.
- Tomasello, Michael. 2005. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge Massachusetts, and London, England: Harvard University Press.
- Tuomela, Raimo. 2005. "We-Intentions revisited." *Philosophical Studies* 125: 327-69.
- Tuomela, Raimo. 2020. "Non-Reductive Views of Shared Intention." In *The Routledge Handbook of Collective Intentionality*, edited by Kirk Ludwig Marija Jankovic. London and New York: Routledge.
- Valentina Fantasia, Hanne de Jaegher and Alessandra Fasulo. 2014. "We can work it out: an enactive look at cooperation." *Frontiers in Psychology* 5.