

**UCLA**

**UCLA Electronic Theses and Dissertations**

**Title**

Joint Action, Speech, and Asymmetric Relations

**Permalink**

<https://escholarship.org/uc/item/7dx378kj>

**Author**

Samuel, Ayana Theresa

**Publication Date**

2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Joint Action, Speech, and Asymmetric Relations

A dissertation submitted in partial satisfaction of the  
requirements for the degree Doctor of Philosophy  
in Philosophy

by

Ayana Theresa Samuel

2022

© Copyright by

Ayana Theresa Samuel

2022

# ABSTRACT OF THE DISSERTATION

Joint Action, Speech, and Asymmetric Relations

by

Ayana Theresa Samuel

Doctor of Philosophy in Philosophy

University of California, 2022

Professor Pamela Hieronymi, Co-Chair

Professor Alexander Jacob Julius, Co-Chair

Everyday life is full of situations in which we act not only as individuals, but also with others, as parts of informal groups or structured collectives. For example, a group of friends spend a day hiking a trail together, a guardian offers a hand to stabilize a young child's tentative first steps, or an academic department draws up policies that facilitate student instruction. Philosophers call this category of action "joint action," and agree that it is an important and distinctive way in which we exercise our agency. However, there has been a tendency in the philosophical literature to explain the phenomenon by analyzing highly structured and cooperative cases of joint action like two or more agents painting a house together or moving a

heavy piece of furniture together. On these analyses, two or more agents are engaged in joint, rather than individual action when they jointly and intentionally pursue goals that they have in common, all while abiding by principles of rationality which demand that they act in ways that are both conducive to achieving the shared goal and responsive to the actions of their co-participants.

In this dissertation, I argue that these standard analyses of joint action do not provide us with a comprehensive understanding of the phenomenon, because they ignore important cases of joint action that do not fit into the highly structured and cooperative paradigm. To address this gap in the literature and bring us closer to a comprehensive understanding of joint action, my dissertation uses case studies of the speech act of telling and early childhood development to show that joint action also occurs in situations which involve little to no cooperation or strongly shared goals. Instead, as we see when we investigate these atypical cases, joint action is consistent with competition and antagonism, and does not require participants to have the sophisticated theory of mind needed to track and adjust to the goals and actions of their co-participants in a way that coheres with principles of rationality.

The dissertation of Ayana Theresa Samuel is approved.

Joshua David Armstrong

Samuel John Cumming

Barbara Herman

Pamela Hieronymi, Committee Co-Chair

Alexander Jacob Julius, Committee Co-Chair

University of California, Los Angeles

2022

*For my mother, Veronica Samuel*

## TABLE OF CONTENTS

<i>Acknowledgements</i>	vii
<i>Vita</i>	ix
0. Introduction	1
1. Chapter 1: Acting Jointly with Others	6
2. Chapter 2: Telling as Joint Action	34
3. Chapter 3: Telling, Indirection, and Joint Action	71
4. Chapter 4: Developmental Joint Action	103
5. Conclusion	142



## ACKNOWLEDGEMENTS

I am extremely grateful to my committee, who have been instrumental in the process of transforming the amorphous ideas with which I began into the arguments I defend in this dissertation. Pamela Hieronymi has been a generous and patient mentor who has provided me with years of incisive advice that I credit with making me a much better philosopher. AJ Julius, through many conversations under the pine trees outside Luvalle, has taught me to be unafraid of pursuing the difficult and hard to articulate ideas that are no less important for being so. Josh Armstrong has helped me to think through and address several challenging issues that come up in this project. Sam Cumming has an unfailing ability to immediately grasp and improve my views and I am grateful for his patient guidance through literatures outside of my usual area of research. Barbara Herman has pushed me to be bold with my ideas and not to shy away from taking up some of the messier aspects of life in my work.

Many thanks to my friends, old and new, who have in various ways made the process of writing this dissertation easier. My writing partners Jenna Donohue, Amber Kavka-Warren, Piera Maurizio, and Ally Peabody-Smith provided motivation when I needed it and transformed the solitary act of writing into a collaborative one. Rima Basu and Gabby Johnson have on countless occasions provided advice about how to navigate the ins and outs of academic philosophy. I am also grateful for the mental and emotional support provided by Sarah Beach, Rima Hussein, Kim Johnston, Tanya Kostochka, Elli Neufeld, Jackie Perez, Ivan Rangel, Megan Robb, and Andrew Stewart.

Thanks also goes to the wider philosophical community here at UCLA. Among the faculty, I am grateful for conversations over the years with Tyler Burge, Daniela Dover, Calvin Normore, Michael Rescorla, and Sean Walsh. Special thanks to Katie Elliott for being a great

teacher and another major source of support during my time here and for being an especially valuable resource on the job market. I also want to thank Doug Myers, our Student Affairs Officer, for his unending patience and help with countless administrative issues over the years. The graduate student community in the philosophy department at UCLA is welcoming and uncompetitive, and I have benefitted from the many valuable conversations I have had with its members, both past and present. A special thank you goes to members of the Ethics Writing Seminar for many interesting and fruitful discussions over the years.

I would also like to recognize some of my former professors at Ryerson University: David Checkland, David Hunter, Jo Kornegay, my “philosophy mother,” and Bob Murray. I likely would not have undertaken this journey without their encouragement years ago.

To Colleen O’Reilly, who has been a tremendous help in the final and crucial stages of writing this dissertation: thank you for your diligent work in helping to put this document together.

Finally, I want to thank my family. Here in Southern California, Martha, Lans, and Fanta have been a great source of support and have provided me with a home away from home. My mother, Veronica, is many miles away in Toronto, but none of this would have been possible without her.

## VITA

### EDUCATION

- 2019 C.Phil. in Philosophy, University of California, Los Angeles  
2011 – 2013 M.A. in Philosophy, Ryerson University  
2013 (Spring) Visiting Student in the Department of Linguistics and Philosophy, MIT  
2007 – 2011 B.A. (Hons) in Arts and Contemporary Studies – English, Ryerson University

### WORKSHOPS AND CONFERENCES

- 2022 “Telling, Indirection, and Joint Action,” American Philosophical Association: Pacific Division meeting  
2021 “Telling, Indirection, and Joint Action,” European Society for Philosophy and Psychology Conference; Ethics Writing Workshop, UCLA  
2021 “Telling as Joint Action,” Reason, Action, and Mind Speaker Series, Ryerson University; Department of Philosophy Speaker Series, York University  
2020 “Joint Commitment and Joint Action,” Ethics Writing Workshop, UCLA  
2019 “Overttness and Joint Action,” Ethics Writing Workshop, UCLA  
2018 “‘Ought Implies Can’ and Interpersonal Moral Conflict,” Proposition Talk Series, UCLA  
2017 Comments on “Responsibility for Testimonial Injustice,” USC/UCLA Graduate Student Conference

### FELLOWSHIPS AND AWARDS

- 2021 Dissertation Year Fellowship, UCLA  
2021 Distinguished Teaching Award, UCLA (awarded to 5 Teaching Assistants university-wide)  
2021 Departmental Teaching Award, Department of Philosophy, UCLA  
2017 – 2020 Doctoral Fellowship, Social Sciences and Humanities Research Council of Canada  
2017 Graduate Summer Research Mentorship, UCLA (with Pamela Hieronymi)  
2016 Graduate Summer Research Mentorship, UCLA (with Barbara Herman)  
2013 Michael Smith Foreign Studies Supplement, Social Sciences and Humanities Research Council of Canada  
2012 Joseph-Armand Bombardier Master’s Scholarship, Social Sciences and Humanities Research Council of Canada  
2012 Ontario Graduate Scholarship (declined)  
2011 Graduate Award, Ryerson University

## Introduction

In trying to tell a temperamental boss that she would prefer not to work late, a worker might choose to hedge her speech in a way that maintains plausible deniability. She says things like: “Oh, this would be the fourth day in a row” or “My family is beginning to forget what I look like.” These speech acts hint at but do not directly express the fact that the worker wants to ask her boss to go home at the end of normal working hours. Yet, it would not be surprising to us if the boss were to figure out her intended meaning and respond by saying (perhaps): “It’s hard on all of us, but you need to stay late.” In other words, it would not be surprising if the boss were to figure out the implied content of the worker’s speech and make it explicit between them.

My dissertation shows that cases like this one, in which interlocutors retrieve and disambiguate a speaker’s implied but unstated message, involve joint action. This might seem surprising to us. After all, paradigm cases of joint action involve agents who share goals and work in a cooperative, rational, and mutually responsive manner to jointly achieve these goals. For example, two or more agents who endeavor to jointly pursue their shared goal of painting a house together would have to cooperate and be mutually responsive to one another’s sub-plans about paint colors or how to divvy up the task in order to secure their end of painting the house together. However, in the case of the worker and her boss, the aims of the two agents are in direct tension: the worker wants to reduce how much her work cuts into her free time and convey that she would like to go home without committing herself to a request to go home, while her boss wants to increase productivity and deny the speaker’s implied request to go home.

Paradigm cases are misleading because they tend to produce individualist accounts of joint action which reduce the phenomenon to the plans and goals of the individual participants. This is in part because walking together or painting a house together are actions that we *can*

perform as individuals, so it is in some sense natural to analyze multi-person versions of these actions in terms of each individual's plans or goals to perform these actions in concert with others. However, we cannot understand how the worker in our opening case manages to communicate her meaning to her boss if we consider their contributions to the exchange separately. The worker's individual act seems to at best commit her to what Elizabeth Fricker (2012, 79) calls a "fuzzy" set of similar messages about how much she has been working recently, rather than a specific request which her boss denies. Similarly, we cannot treat her speech act as a cooperative joint project that realizes a shared goal. We can make sense of what seems to go on in this case only when we see that despite the apparent antagonism and lack of shared goals, the worker and her boss jointly determine what speech act is performed.

When we see that joint action need not conform to the contours of paradigm cases like painting a house together, we are better placed to appreciate the full scope of the phenomenon. My dissertation presents three atypical cases of joint action that are inconsistent with standard analyses in the literature. I argue that joint action can occur in situations that are antagonistic or characterized by very few shared goals or interests and can also involve participants without the capacity to represent or pursue shared goals.

I begin, in Chapter 1, by surveying the current state of analytic philosophical thought on joint action. Many commentators agree that joint action is distinguished from merely coincident individual action by the intentional states of the agents involved: our going for a walk together is distinct from what happens when I happen to fall into step with a stranger on a path, because in the former case, we both intend to pursue the shared goal of walking together. Reductive accounts of joint action (see Bratman 1999; Tuomela and Miller 1988) model the shared intentions of joint activity on typical individual intentions. But a related form of individualism is

also found in Gilbert's (1990) non-reductive account because she deploys the strategy of beginning with the ostensibly simple case of individual action and trying to determine what would need to be added to transform it into joint action. As critics like Baier (1997) and Stoutland (1997) point out, this individualism is misleading because it steers us away from a full understanding of atypical or undertheorised forms of joint action that resist these individualist and reductive analyses.

The second through fourth chapters of my dissertation provide an analysis of atypical cases of joint action. In Chapter 2, I argue that the speech-act of telling is a form of joint action that resists standard analyses of the phenomenon. To do so I take up a recent argument by John Greco (2020) which uses Michael Bratman's (1999) influential and reductive account of joint action to try to show that acts of telling that successfully transmit knowledge to an interlocutor involve joint agency. I argue that although Greco is right that telling involves joint action, we cannot analyze telling by appeal to Bratman's account of shared intentions. This is in part because speakers and interlocutors in acts of telling do not need to share intentions because they do not need to have the same goals at the start of the exchange in order for an act of telling to take place. In fact, successful acts of telling take place even when the situation between the speaker and the interlocutor is highly antagonistic or hostile. In place of Bratman's account of shared intentions, I suggest that we should instead understand telling as a form of joint action because it involves a speaker and interlocutor in a state of shared understanding as they jointly determine the common ground of their exchange.

In Chapter 3, I apply my argument that telling is a form of joint action to a problem that seems to be raised by indirect forms of speech for telling-based views of testimony. According to telling-based views of testimony, paradigmatic acts of telling involve a speaker who publicly

takes responsibility for the truth of some proposition  $p$  by way of her overt statement that  $p$ . On these views, the speaker's overt statement and consequent assumption of responsibility is what licenses an interlocutor's belief in what a speaker says—it is why testimonial knowledge is justified. In indirect speech, however, speakers do not openly state that  $p$ , and this seems to allow speakers to maintain plausible deniability and avoid responsibility for their speech, which in turn blocks the transmission of testimonial knowledge. However, I suggest that this is an unacceptable conclusion. Because indirect speech is very common, this result would undermine much of our testimonial knowledge. To avoid this conclusion, I highlight some cases in which interlocutors seem to contribute to the determination of ambiguous speech acts and thereby prevent the speaker from avoiding responsibility for what she says. I argue that we can make sense of this only if we understand telling as a form of joint action.

In Chapter 4, I argue that the reciprocal interactions that take place between infants under the age of nine months and their adult caretakers constitute joint action. This form of what I call *developmental* joint action cannot be explained by the paradigm of joint activity, in part because infants lack the ability to engage in sophisticated intentional behavior required by standard accounts. But infants under nine months also cannot meet the requirements of more minimal accounts of joint action that were recently put forward by philosophers who noticed the standard accounts' apparent exclusion of young children. This is because the children-accommodating accounts of joint action retain the *triadic* structure of standard examples in the literature in which two or more people pursue an external goal that they happen to share. But because infants do not develop the ability to engage in triadic behaviors until they are about twelve to eighteen months old, these accounts also exclude joint action in infants. To explain why these reciprocal interactions should nevertheless be thought of as joint action, I develop an account that explains

how the *dyadic* structure of the reciprocal interactions that take place between the infant and her primary adult caretaker can nevertheless be understood as a form of joint action.

In sum, philosophical treatment of joint action has been fairly narrowly focused on highly cooperative and structured actions like walking together and painting a house together. My dissertation aims to show that there are potentially many other forms of joint action that diverge from this basic structure. My hope is that when we understand the full scope of the phenomenon, we will be better placed to understand some of the important moral issues that arise when we act jointly with one another. Our dealings with others, after all, naturally generate opportunities to create new moral relationships and infringe on moral boundaries. In our worker-boss case, for example, we see how a worker and her boss might jointly produce an outcome that the worker, who holds less power in the relationship, wants to avoid. This paves the way for future research into how it is that the speech of marginalized persons may be transformed against their will, and why they might exert less control over their speech than those with more power.



## Chapter 1: Acting Jointly with Others

### 0. Introduction

Everyday life is full of situations in which we act and exercise our agency not only as individuals, but also with others, as members of informal groups or structured collectives. For example, a group of friends spend a day hiking a trail together, a guardian offers a hand to stabilize a young child's tentative first steps, or a school board makes decisions and draws up policies. Philosophers have long recognized this distinction between the actions of individual agents and the actions of formal or informal groups of agents, but there is significant disagreement about how it should be characterized. In what sense do the agents in the cases above act with each other? And how is this form of action distinct from or related to individual action?

A first step towards answering these questions is found in John Searle's observation that some individual actions are outwardly indistinguishable from what I will refer to as "joint action" (Searle 1990; Roth 2017). Consider, for example, a group of students who walk down a particular path on campus in between classes, and another group of students who march down the same path in order to protest or bring attention to some issue. If at least some of the students in the latter case are not holding signs, chanting, or otherwise signaling their membership in the group of protestors, then some of the students in the latter case will be outwardly indistinguishable from students in the former case who are merely walking down the path to get to their destination. This point is also nicely illustrated by "flash mobs," who exploit the fact that they can appear to be individuals who, for example, happen to be taking the train at the same

time, but are in fact members of an orchestra who jointly take the train and plan to perform a piece of music.<sup>1</sup>

It has become standard practice in the joint action literature to distinguish these collections of outwardly similar individual actions from genuine joint action by reference to the intentions of the agents involved.<sup>2</sup> The thought is that while agents engaged in individual actions will hold the usual, individual intentions to perform the action in question, groups of agents engaged in joint action will hold intentions that in some way refer to the collective structure of the action. This move is a plausible one, particularly if we accept Donald Davidson's claim that actions are behaviours that are intentional under some description, because it would seem to follow that joint actions are intentional under some description (Davidson 1980).

What does it mean to say that actions are behaviours that are intentional under some description? In the first part of the claim, Davidson suggests that actions are not mere bodily movements, but bodily movements that can be explained by the agent's intentions.<sup>3</sup> To use Davidson's example, when I raise my arm and flip the light switch, that behavior is explained by my intention to turn on the light. In the second part of the claim, Davidson clarifies that actions

---

<sup>1</sup> Flash mob in the Copenhagen metro: [https://www.youtube.com/watch?v=gww9\\_S4PNV0](https://www.youtube.com/watch?v=gww9_S4PNV0).

<sup>2</sup> As I later discuss, an exception to this view is found in some minimal accounts of joint action that attempt to replace intentions with less cognitively demanding states, see e.g., Tollefsen (2005) and Butterfill (2012). Another opposing view is found in Chant (2007). Chant argues that "unintentional collective action" is an important subset of collective action, one that any theory of collective action should be able to accommodate. This does not necessarily pose a problem for the general claim that agents' intentions are an important distinguishing feature of joint action, because collective action is generally thought to be a wider category than joint action. See e.g., Chant (2007) and Hammond (2016). Thus, the idea that there is unintentional collective action is consistent with the idea that there is a subset of actions—joint actions—which do characteristically involve agents who intend to perform some action together.

<sup>3</sup> I exclude many of the details, but here I refer to Davidson's later view of intentional action, according to which an agent's pro-attitudes or desires and beliefs cause the intention which in turn causes the action. See Davidson (1980).

are intentional under *some* description, because one and the same action can often be intentional under one description and unintentional under another. My act of turning on a light switch may alert a burglar who is lurking outside, however, that action is not intentional under this latter description because I did not intentionally flip the switch in order to alert the burglar. Many philosophers accept this basic view of (individual) action as intentional behavior, and this has made intention a natural starting point for analyses of joint action.<sup>4</sup> To return to our earlier example, on this basic view, we would distinguish between the actions of the individual students and the action of the protestors by referencing the fact that the protestors' action is intentional under the description "marching to protest *x* issue," while the actions of the individual students that happen to be walking on the same path are not.

In this chapter, I survey the current state of analytic philosophical thought on joint action, which has largely focused on describing the nature of the intentions held by participants in joint action, including their structure and how they come about. For example, reductive and non-reductive accounts of joint action disagree about whether the intentions characteristic of joint action—which I will refer to as shared intentions—can be understood in terms of individual intentions or whether shared intentions are in some way irreducible. Notable figures in this debate include Michael Bratman (1999), who presents the most sustained and influential defense of this reductive understanding of shared intention.<sup>5</sup> For Bratman, shared intentions are interlocking complexes of individual intentions that help participants coordinate and plan their shared activity. Margaret Gilbert (2009) famously rejects Bratman's broadly reductive strategy

---

<sup>4</sup> There is, however, significant disagreement about the nature of intention, but the points in this chapter can be made without getting into that debate.

<sup>5</sup> See also Kutz (2000), McMahan (2001), and Miller (2001).

and its exclusion of the mutual obligations that shared intentions seem to give rise to. Instead, Gilbert defends the view that two or more agents share an intention when they jointly commit to intend “as a body” or “plural subject” to perform some action  $x$ .<sup>6</sup>

But others (Baier 1997; Stoutland 1997) have found this debate unsatisfactory, arguing that the joint action literature in general suffers from an individualist bias. They point out that despite nominally rejecting the individualist, reductive strategy, even anti-reductionists like Gilbert and Searle retain individualist elements in their accounts. At least part of the reason for this bias is methodological. Contemporary philosophers of action have generally approached the phenomenon of joint action through analysis of examples like walking together (Gilbert 1990), painting a house together (Bratman 1999), and lifting a heavy piece of furniture together (Tuomela 2005). These examples feature actions that can all in principle be performed by individuals, and as Annette Baier (1997, 43) observes, this has led to analyses that cast joint action as a joining of individual actions. However, for Baier, this sits uneasily with the following data: first, that some actions (communicating, contracting, marrying) can only be done with others, and second, that our first experiences of even paradigmatically individual actions like reasoning or reading are communal rather than individual (1997, 26-31).

Several recent accounts of joint action have seemingly taken up this latter point of Baier’s, arguing that much of the literature ignores the reality of joint action in young children (Tollefsen 2005; Butterfill 2012; Pacherie 2013; Story 2021). According to these accounts, this is because standard analyses of joint action build in conditions for sharing intentions that are beyond the cognitive capacities of young children, especially those who have not yet developed a robust theory of mind (Wellman 2002). In response, these alternative, children-accommodating

---

<sup>6</sup> Searle (1990) also defends a non-reductive account of shared intention.

accounts of joint action attempt to weaken the cognitive requirements of standard accounts of shared intention.

In what follows, I first clarify the target phenomenon of joint action by making note of a few distinctions. Then, I provide an overview of the debates mentioned above, beginning with the classic debate between reductive and non-reductive accounts of joint action. I then turn to the small number of philosophers who have criticized both sides of this debate on the grounds that there is an individualist bias in their methodology. Next, I discuss some recent attempts to expand the domain of joint action to include joint action in young children. As I try to highlight throughout my discussion, there are several gaps in the literature on joint action that ought to be addressed if we are to gain a complete understanding of this phenomenon. While recent accounts of joint action in young children are an exception, the literature is largely focused on simple, often pre-planned and highly cooperative activities, to the exclusion of cases of joint action that involve antagonism or imbalances in the contributions of participants.

### **1. What is joint action?**

Before turning to some of the debates that have shaped current philosophical discussion of joint action, it is useful to first clarify the target notion by making a few distinctions. First, within the broad category of actions that we perform with others, it is possible to distinguish between the actions we undertake as informal groups and actions we undertake as structured, institutional collectives. In the informal case, participating agents involved are often referred to as a plurality (Stoutland 1997, 46-7). Thus, we might say of a group of friends that *they* play a board game, watch a movie, paint a house, etc. More formal collectives, however, are generally referred to as singular, rather than plural entities. Referring to a company, we might say that its

profits are up in the second quarter, or it has decided to lay off workers to increase profits for shareholders, and so on. Notably, informal groups also seem to come into being more easily than formal institutions or organizations, while formal groups tend to have a history that often precedes and can outlive any of its individual members (47-8). Formal or institutional collectives raise important issues about the relationship between the individual and the collective, and particularly the relationship between individual and collective responsibility. However, the joint action literature has, with a few exceptions, focused on the joint actions of informal groups of usually no more than two people.<sup>7</sup>

In addition to the distinctions between the different kinds of groups that may engage in joint action, it is also possible to distinguish two broad act-types discussed in the joint action literature. The first act-type is what we might call *necessarily* social, because these actions cannot be performed by individuals, but instead require the contributions of at least two people.<sup>8</sup> Examples of this category of action includes conversing, buying, selling, contracting, and marrying. Thomas Reid calls this category “social acts of mind” and defines them as actions that “necessarily imply social intercourse with some other intelligent being who bears a part in them” (1788, 537).

But because our lives are deeply interrelated, Reid’s brief definition potentially captures actions that philosophers do not typically regard as joint action. As Reid’s contemporary David Hume observes, “the mutual dependence of men is so great, in all societies, that scarce any human action is entirely complete in itself, or is performed without some reference to the actions

---

<sup>7</sup> Exceptions to the general focus on informal, two-person groups are found in McMahon (2001) and Kutz (2000).

<sup>8</sup> I use “contribution” in its broadest sense here and do not assume that it involves any outwardly observable action.

of others” (Hume 1748, Section VIII, Part I). Hume’s point here is that even those actions that we ordinarily take ourselves to perform as individuals refer to or rely on the actions of others. Going for a solo walk in a nearby park, for example, is an action that is made possible by the actions of many others: from the municipal council members, landscapers, and gardeners who create and maintain the park to the other people in the park who understand that they must behave in certain ways to facilitate our shared use of the park. However, taking a solo walk in the park is not the kind of action that philosophers are typically concerned with when they think about and analyze joint action.

So, the basic idea that some actions “necessarily imply social intercourse” with another agent for their completion cannot by itself help us identify the first broad act-type of necessarily social actions. As Hume suggests, this is because our lives are deeply interrelated when we live together in societies: everything we do potentially “bears a part” in the actions of others, even if all we do is refrain from interfering in their actions. But this issue can be dissolved if we build a little more into Reid’s definition. Because while actions like conversing, buying, or contracting cannot take place without the contributions of others, this is not all that is distinctive about them. In addition, people involved in necessarily social actions like conversing or contracting are in a state of mutual awareness of what they do together. This excludes cases like taking a solo walk in the park, because although both the municipal council members and the residents under their jurisdiction will likely be aware that the council’s creation and maintenance of the park makes the residents’ use of the park possible, this awareness is not mutual.<sup>9</sup> Let us therefore say that the

---

<sup>9</sup> Mutual awareness can be thought of here as an iterative state: we are mutually aware of some x when I know that you know that x, and you know that I know that you know that x, and so on. This is unlikely to describe the situation between the resident on a solo walk and municipal council members because the resident and council are more likely to be separately aware of the part the council plays in facilitating the solo walk.

first act-type is necessarily social because these actions are interdependent and involve two or more people in a state of mutual awareness of what they do together.

The second broad act-type within joint action is neutral with respect to the jointness or social character of the action because it includes actions that can be performed jointly with others but can also be performed by individuals. This includes actions like painting a house, moving a couch, or going on a trip. While informal groups of two or more agents can paint houses, move pianos, or go on trips, so can individual agents. Crucially, the tendency in the joint action literature has been to focus on these neutral act-types over social act-types. Bratman, who draws exclusively on neutral act-types in his reductive account, defends this strategy by claiming that analysis of social act-types risks a problematic circularity, because social act-types build the shared intentions that we aim to understand into the action we use to understand it (Bratman 1999, 114).<sup>10</sup>

In sum, joint actions may be performed by either informal groups or formal collectives, though the literature on joint action has tended to focus on the former. There are also two broad act-types that fall into the category of joint action: necessarily social act-types like conversing, promising, or contracting, as well as neutral act-types like painting a house or taking a trip. Of the two act-types, the literature on joint action has overwhelmingly focused on neutral act-types. Thus, the literature has tended to focus on analyzing neutral act-types that are performed by informal groups of usually no more than two people. In Section 3, I turn to accounts that criticize this general trend in the literature on the grounds that analyses of two people who jointly perform one of these neutral act-types are too narrowly focused and guilty of an individualist bias. But

---

<sup>10</sup> Tollefsen follows Bratman in focusing on neutral act-types in her account of joint action in children (2005, 77)



first, I turn to the debate about the nature of shared intention which has occupied much of the joint action literature.

## **2. Shared intention: reductive and non-reductive views**

As we have seen, many accept that shared intention is the characteristic feature of joint action. This is suggested by the fact that if asked what they are doing, participants in a joint action would likely reply that “we are *x*-ing” which seems to express an intention with a shared content (Searle 1990, 402-3). For example, if Jack and Sue are taking a walk together and bump into a friend who asks what they are up to, Sue (or Jack) might reply by saying “We’re taking a walk” in the same way that Sue might reply “I’m taking a walk” if she happened to be taking a solo walk that day. If we take each reply as an expression of the intention that Sue happens to hold in the two scenarios, then it seems to follow that Sue’s intention when she walks with Jack is distinct from her intention when she walks alone. Specifically, Sue seems to express a shared intention to walk with Jack when Sue and Jack walk together.

How should these shared intentions be understood? There are two broad approaches to this question. The first, reductive approach, analyzes shared intentions in terms of the individual intentions of participants, while the second, non-reductive approach holds that the shared intention is basic, and not reducible to individual intentions. I begin with an overview of the reductive approach.

There are two related thoughts that seem to motivate the broadly reductive explanation of joint action. The first is that agents can only intend what they take to be within their control, which Baier has called “the limited sovereignty of intentions” (1997, 25). The second is that agents only have the requisite sort of control over their own decisions and actions. As Velleman

(1997) explains, this raises a problem for the very concept of shared intention because there is a tension between the idea that we can only intend what we take to be within our control and the idea that joint activity seems to require that we in some sense share intentions with someone else—someone whose decisions and actions are not (or at least may not be) under our control.

Because of this, several reductive accounts limit the content of participant intentions to the role that the participant plays in the joint action (Tuomela and Miller 1988; Kutz 2000; McMahon 2001). To put it another way, on these accounts, the shared intention is in place only in a case where each participant in the joint action intends to (1) *do her part* in the joint action while believing that (2) the other participants will also do their part and (3) that all participants believe that all participants will do their part (Tuomela and Miller 1988, 375). This move plausibly avoids the problem of the limited sovereignty of intentions. However, we might wonder whether participatory intentions and expectations about the contributions of other participants is too minimal. After all, coordinated action among strangers in public spaces (of the kind that occurs on a bus or train) could meet these conditions. To address this worry, some proponents of this view point out that in addition to an intention to do one's part in a joint action, one also plausibly intends to act, as a group, in the pursuit of the joint action (Tuomela and Miller 1988, 376; McMahon 2001, 289-290). Others have simply denied that there is a problem with holding that participants act only on participatory, rather than group intentions (Kutz 2000, 22).

In his highly influential reductive account, Bratman finds a different way around the problem of intending actions outside of one's control. According to Bratman, we intend to perform some action joint action J if (1) we both hold the relevant individual intentions (expressed as "I intend that we J"), (2) each of our intentions is "in accordance with and because

of' the other's intention to act with us and our subplans (that is, how we plan to achieve our shared goal) are consistent, and (3) all of this is common knowledge between us (Bratman 1992, 334-5). Bratman also thinks that the meshing of the participants' subplans is a "rational requirement" which can be traced to the norms of consistency and coherence that guide our individual intentions.<sup>11</sup> How do these conditions avoid having each participant intend what she does not control? For Bratman, it is possible to simply build some measure of control or influence into the conditions of the individual intention. Thus, an intention of the form "I intend that we J" could only be held by someone who sees the actions of the other participants as being at least partially up to her (1999, 116).

In contrast to the broadly reductive views described above, Velleman defends the claim that it is possible to "literally [share] a single intention" (1997, 29). Velleman frames his thesis in this way to highlight the fact that reductive accounts do not really seem to explain how it is that intentions are shared. Instead, these accounts suggest that individual intentions can be formulated in a way that amounts to shared intention (33). Bratman, for example, acknowledges that in his account, shared intentions are not intentions at all but "states of affairs consisting primarily of appropriate attitudes of each individual participant and their relations" (1999, 111).

Velleman begins his analysis with the view that ordinary individual intentions are the "attitudes that resolve deliberative questions, thereby settling issues that are up to you" (1997, 32). This includes, for example, the question of whether to go for a walk, cook a meal, or read a book. In each of these cases, I may resolve and thereby settle the deliberative question of, for

---

<sup>11</sup> Roughly, Bratman's thought is that given my intention to perform some action with you, rational norms of consistency and coherence will require me to ensure that my plans to perform this action are consistent with yours.

example, whether I should go for a walk by forming an intention to go for a walk, because going for a walk is an issue that is “up to me.”

Given this view of individual intention, an intention is shared only in cases where multiple people resolve a deliberative question that is somehow up to them. But here again, the problem of intending that which is not under one’s control resurfaces. As Velleman puts it, the “logical space of decision making is open only to those who are in a position to resolve an issue, and it admits of only one resolution per issue” (35). That is, whether we go for a walk must be up to us, but it seems that this deliberative question can only be resolved once. Thus, it seems that if you make a decision about the deliberative question of whether we will go for a walk, then you have resolved the issue and there is nothing left for me to decide: your forming the intention that “we” are going to go for a walk seems to entail settling my going for a walk with you, as well as your going for a walk with me. This would essentially close me off from an intention that has to be partly up to me if it is to count as a shared intention.

Velleman ultimately tries to solve the problem of the limited sovereignty of intentions by drawing on an idea from Gilbert (1990). Velleman interprets Gilbert as proposing that several individuals can exercise conditional discretion over some issue “in such a way that their conditionally settling the issue separately adds up to their categorically settling it together” (1997, 36). For Velleman, each individual must make their settling of the deliberative question public, perhaps by expressing it in speech. For example, Sue and Jack may together settle the question of whether they will take a walk together when Sue says, “I will if you will” and Jack says “Then I will.” In this way, Velleman argues, it is possible for these agents to exercise discretion over a single issue, to have their individual intentions combine to form a single intention that governs the behavior of both.

Searle (1990) also proposes an account according to which “we intend to *x*” expresses a genuinely shared intention. For Searle, this intention is primitive, and not reducible to individual intentions of the form “I intend that we *x*” and beliefs about whether the participants in the joint action will do their part. He defends his position by constructing a counterexample to reductive accounts of shared intention in which a graduating class of business students each form individual intentions to selfishly pursue their own interests as a way of realizing Adam Smith’s theory of the invisible hand (1990, 404). According to Searle, if we also add to the story that each graduate believes that each of the other graduates has also formed this intention because they were all similarly educated, they would meet the conditions of reductive accounts of shared intention.<sup>12</sup> That is, they would each (1) intend to do their part in advancing the greater good by pursuing their own selfish interests, (2) believe that the other graduates will do the same, and (3) believe that every other graduate also believes that each graduate will do their part in advancing the greater good by selfishly pursuing their own interests (Searle 1990, 405). However, the business school graduates are intuitively not engaged in joint action, thus for Searle, reductive accounts of joint action are mistaken.

Notably, however, Searle believes that all intentions, including his irreducibly shared intentions, can only be held by individuals. Thus, although the content of the intention is irreducibly shared, the shared intention itself is not literally shared in any sense among the participants. In fact, in Searle’s view, a brain in a vat could hold such an intention, even though it would be wildly mistaken about the outside world and the others with whom it intends to perform some *x* (1990, 407).

---

<sup>12</sup> This would not be a counterexample to Bratman’s account, however, because the business school graduates would not intend to mesh their subplans with the subplans of their fellow graduates.

For Gilbert (1990), who is a frequent and vocal critic of the reductive strategy, an important part of joint action is the commitments, obligations, or entitlements that seem to attend joint action. In Gilbert's account, Sue and Jack are engaged in the joint action of taking a walk together when they mutually express their readiness to be jointly committed to the task of going for a walk together. Jack and Sue form this joint commitment when they express a "willingness to constitute with the other a plural subject of the goal" that they undertake some activity together. In turn, a "plural subject" is formed when each or two or more persons offers up their will to be part of a "pool of wills" which is dedicated "as one" to a goal. These wills are bound or *committed*, according to Gilbert, "simultaneously and interdependently" (1990, 7). For Gilbert, this makes joint commitment unlike promising because each person does not unilaterally bind herself to the goal in question, hoping that the other follows suit, nor does one person say that the other may regard her as committed once he has committed himself. Instead, Gilbert thinks that each person conditionally binds herself, such that only when everyone is committed is *anyone* committed.

The plural subject that is formed when Jack and Sue commit to the goal of going for a walk together is irreducible, and one of the ways Gilbert tries to establish the irreducibility of the plural subject is by suggesting that it has its own form of rationality. For example, if Sue accepts the premise that she and Jack have the goal of walking down the street together, and it is also true that "Jack is drawing ahead," and she believes that the best way she can help achieve the shared goal is to tell Jack to slow down, then it seems that Sue ought to tell Jack to slow down. For Gilbert, the existence of these obligations also works against reductive accounts like Bratman's, because it seems that we can explain these obligations without reference to Sue's personal goals or desires. In a similar vein, Antonie Meijers (2003) argues that these claims and

obligations cannot be adequately explained by a reductive account like Bratman's, because my individual intention to go for a walk with you does not seem to give rise to my having an obligation to you to keep pace, or to avoid veering off the path, and so on. For Meijers, because the intentions of Bratman's account are individual intentions, it seems that they can at best give rise to obligations to oneself.

An outcome of the irreducibility of the plural subject is that no individual participant can modify the shared intention without the approval of all the other participants. Both (or all) participants are obligated to do their part in realizing their shared goal until they are released from these obligations by the other participant(s) or they accomplish the shared goal.

For Gilbert then, participants in joint activity are obligated to one another to do their part in the joint activity. These obligations are generated when participants jointly commit to do their part in the joint activity. That is, when they form a plural subject that is dedicated to undertaking some shared goal. In sum, Gilbert's view is that two or more people act together if:

- 1) Jack intentionally expresses to Sue his willingness jointly to accept with Sue that Jack is to walk with her down a particular path.
- 2) Sue did the same.
- 3) The fulfillment of the first two conditions is common knowledge between Jack and Sue. (Gilbert 1996, 296)

In sum, a major debate in the current joint action literature has been focused on the question of whether the shared intentions that are characteristic of joint action are basic or can be reduced to ordinary, individual intentions. While several (Bratman 1999; Tuomela and Miller 1988; McMahon 2001; Kutz 2000) have opted for the latter view because it seems to more easily avoid the problem of the limited sovereignty of intention—the idea that we cannot intend what is

not under our control—others, like Velleman, have developed accounts that avoid both reduction and the limited sovereignty problem.

Other non-reductive accounts like Gilbert's have emphasised the idea that participants in joint action seem to incur obligations to do their part in the joint action, and that these obligations, for various reasons, cannot be explained by individual intentions. Instead, what is needed to explain these obligations is an intention that is genuinely shared among participants. This is because only shared intentions, unlike individual intentions, can explain why agents who share an intention seem to incur various obligations to each other. Notably, however, these reductive and non-reductive accounts are not as different as they might initially appear. For while proponents of non-reductive accounts like Gilbert avoid analyzing joint action in terms of the individual intentions of participants, there is often a weaker kind of individualism in their accounts.

In the following sections, I turn to some criticisms of the individualism that seems to pervade both reductive and non-reductive accounts of joint action. I begin in the next section with some concerns that have been raised by Annette Baier (1997) and Frederik Stoutland (1997). I then turn to more recent criticisms of the standard accounts' apparent exclusion of young children.

### **3. Individualist bias in joint action**

In similar papers, Baier (1997) and Stoutland (1997) criticize several of the accounts of shared intention described above on the grounds that they display an unwarranted bias towards individualism even as they attempt to explain an important social phenomenon. Baier, for example, argues that despite Gilbert and Searle's open rejection of the broadly individualist spirit



of reductive accounts, both philosophers are guilty of some of the same individualist bias that is found in reductive accounts (1997, 17). As we saw in the last section, Searle, despite arguing that shared intentions are not reducible to individual intentions plus mutual belief, believes that “society consists of nothing but individuals” and that the shared intention in question is not “the product of some mysterious group mind” (1990, 406). This leads Searle to accept that a brain in a vat who is entirely mistaken about the external world and the existence of other individuals could nevertheless possess a shared intention. For Stoutland, this suspicion of “group minds” or social agents, as he calls them, shows that there is a deep anti-social bias “even among those sympathetic to social action” (1997, 52).

A similar suspicion of so-called group entities is found even in Gilbert’s discussion of plural subjects. Baier notes that Gilbert describes her account as “weak analytic individualism” because she believes that the individuals who constitute plural subjects are basic (1997, 17). That is, Gilbert believes that the concept of an individual person cannot be analyzed in terms of a plurality that is itself unanalyzable in terms of individuals (1989, 435-36). Gilbert and others assume, without argument, that it is the individual that is prior and basic, and that our understanding of plural subjects is made possible by looking to the contributions of individual members who pool their wills to then act as a plural subject. Similar assumptions are made by both Searle and Bratman. That is, while Searle does not reduce shared intentions to individual intentions of the form “I intend that we  $x$ ” there is no functional difference between an agent with a Searlian shared intention in mind and an agent with a Bratmanian shared intention in mind (Baier 1997, 21).

In addition to assuming that individual forms of action are more basic or simple and can therefore be used to explain the more complex form of joint action, philosophers have tended to

focus on particular kinds of actions in their accounts. As we saw earlier, the tendency has been to focus on neutral act-types like painting a house, lifting a couch, or taking a trip, which can all in principle be performed by individuals. Thus, the examples deployed by philosophers of action seem to play into the individualist bias by allowing philosophers to analyze joint action and shared intention by starting with the individual form of the action and then trying to determine what needs to be added to the basic, individual form of the action to make it a joint action.

However, as Stoutland observes, this methodology is potentially misleading. This is because, when groups of agents engage in any form of joint action, they do so *as* a group, and, in that respect, “the act is *indivisible* among individual agents” (1997, 48; Stoutland’s emphasis). That is, the action cannot be carved up into parts that correspond to the individual actions of the participants. Stoutland points out that the indivisibility of joint action can be obscured by the kinds of cases we choose to consider: because individual agents may go on walks, we may be tempted to think of walking together as a case of two agents performing a particular act type of walking, and doing so together (49).

But this way of conceiving of joint action seems less neatly applicable to other kinds of cases of joint action, like performing a duet, waging war, or communicating. This is because a single agent cannot perform a duet or wage a war, and so it seems that we cannot understand these agents as each performing a neutral act-type, such as arguing, and doing so together. For Stoutland, a similar point can be made about neutral act-types like painting a house or taking a trip when they are performed as a joint action (49). Because although these actions are sometimes performed by individuals, they are also performed by groups, and when they are, they should be treated in the same way that we treat social act-types (which are always performed by groups). And, because both social and neutral act-types of joint action are performed by groups

of agents, they should be thought of as a single act (e.g., arguing or painting a house together) that is carried out by the group agent.

Thus, on this view, we have reason to be skeptical of Bratman's individualist account which relies on exactly this strategy of analyzing joint actions as if they are all comprised of constituent, individual action. While this strategy might yield plausible analyses of these neutral act-types, it is less clear how Bratman's account would fare in analyzing social act-types like buying or contracting.

A final methodological criticism made by Baier is related to the earlier point that standard accounts of joint action in the literature incorporate, without argument, the idea that individual action is more fundamental or basic than joint action. Baier points out that this ignores the developmental fact that children gradually become more independent and less reliant on their caretakers as they grow older (1997, 29-30). Prior to and during this gradual independence, the child relies on her caretakers when she eats, walks, or begins to communicate with language. At this stage of reliance, the child *acts with* those with whom she finds herself surrounded. For example, the child acts with her caretakers when she uses their limbs or the push toys they have placed in her vicinity to begin to balance on two feet. She also acts with them when she begins to sound out parts of the words and phrases she hears most often and her caretakers complete her attempts by repeating the word in full and demonstrating their understanding of what she aims to say. Very gradually, she learns to do these things on her own, but this is only after she does them with others. As Baier points out, "individual action was something we all had to learn, and [...] we learned it as a departure from common action" (1997, 29).

In the next section, I turn to several recent accounts of joint action that take up the issue of joint action in children. These accounts were developed in response to a tension that seems to

arise when we consider the following two data: first, that young children seem to readily engage in joint action, and second, that standard accounts of joint action build in conditions that require cognitive capacities generally not available to children.

#### **4. Joint action in young children**

Recently, several philosophers have generated a problem for leading accounts of joint action in the following way: leading accounts of joint action such as Bratman (1999) and Gilbert (2009) (though the accounts discussed below all focus on Bratman) often seem paint a picture of joint action as only taking place between equal and cooperating adults who form complex intentions about their joint pursuit of a shared goal. Gilbert, for example, makes this claim fairly explicitly in a discussion of “social groups,” a term which may be applied to two people while they engage in a joint activity like taking a walk together, but not to infants or the profoundly intellectually disabled:

“...it is possible that some so-called families are neither paradigmatic families nor paradigmatic social groups. A case in point might be a single adult with an infant child. While the child is very small, it may be insufficiently developed to enter into the paradigmatic member relationship. The same goes for a single adult who is caring for a severely mentally disabled child, sibling, or parent. With respect to the membership relationship, the condition of autism in a child may be especially pertinent. Such families may be sufficiently like paradigmatic families that they have come to be referred to by the same label. Yet they may be neither paradigmatic families nor paradigmatic social groups.” (Gilbert 2006, 94)

However, young children do seem to engage to in joint actions like building a tower and rolling a ball back and forth. In fact, some psychologists believe that these early joint actions are responsible for later cognitive development (Moll and Tomasello 2007). In response to this problem, Tollefsen (2005), Butterfill (2012), Pacherie (2013), and Story (2021) have advanced alternative accounts of joint action that can accommodate the reality of joint action in young children.

Tollefsen (2005) argues that because Bratman's account of shared intention requires participants to be mutually responsive to and have common knowledge of the intentions of their co-participants in order to effectively coordinate their actions, his account rules out the possibility of joint action in young children. Specifically, for Tollefsen, because mutual responsiveness and common knowledge require a robust theory of mind that children do not develop before the age of four or five years, Bratman's account rules out the possibility of joint action in children younger than that (2005, 81). However, research in developmental psychology suggests that children are readily able to engage in simple joint actions like having a pretend tea party or building a tower beginning between twelve and eighteen months of age.

To accommodate these early cases of joint action in young children, Tollefsen modifies Bratman's account by replacing the mutual responsiveness and common knowledge conditions with minimal, child-friendly versions of these conditions. In place of the requirement to be mutually responsive to the intentions of one's co-participants, Tollefsen suggests a requirement to be responsive to their intentions-in-action, which are less cognitively demanding for young children because they require only an ability to understand the broad aim or goal of an (outward) behavior or action (53). In place of the common knowledge requirement, Tollefsen proposes a

joint attention requirement, on the grounds that children develop this ability by nine months old (84).

Stephen Butterfill (2012) and Elisabeth Pacherie (2013) similarly aim to relax the cognitive requirements of accounts like Bratman's by proposing alternative conditions that could serve similar functions in joint action. Butterfill claims that "there is a simple account of joint action which is compatible with the developmental premise. On the simple account, joint action involves sharing goals and sharing goals requires only an understanding of goal-directed actions and their common effects" (2012, 47). Butterfill therefore does away entirely with shared intention and instead argues that joint action in children requires only the ability to track and share goals.

Pacherie (2013) addresses the problem by drawing on Bacharach's theory of team reasoning. She focuses on children around two years old and proposes a minimal theory of shared intention called "Lite Shared Intention." For Pacherie, children share "lite" intention if they can conceive of themselves as a member of a team that consists of themselves and another and can reason that some action A is the best choice of action for that team (2013, 1833).

In contrast to the views above, Daniel Story (2021) does not directly attempt to relax the cognitive requirements of accounts like Bratman's. Instead, drawing on Abraham Roth's (2014) concept of practical intimacy, or the idea that it is possible to "share an intention issued from a single deliberative source," Story argues that it is possible for the child to act directly on the shared intention issued by an adult partner (2021, 5010). The psychological sophistication required by Bratman's account is essentially, in Story's view, off-loaded onto the adult in the adult-child pair, while the child is able to "freeload" off of the adult's cognitive capacities while being an active participant in the joint action.

These accounts provide important contributions to the literature on joint action which has, in its short history, been narrowly focused on the nature of the internal attitudes held by adult participants in strictly cooperative activities. As these accounts suggest, joint action is not only analyzable in terms of shared intention but potentially a whole host of other attitudes. This helps to highlight and explain forms of joint action that the current literature has largely shied away from: for example, brief or transient joint action which does not involve forethought or planning, joint action which involves imbalances in the capacities or contributions of participants, and joint action that is not cooperative under some description.

## **5. Conclusion**

In this chapter, I have provided an overview of the current philosophical literature on joint action. I began by clarifying the target notion of joint action by making note of a few key distinctions. Then, I discussed the debate about the nature of the shared intentions that many accept as the characteristic feature of joint action. For commentators like Bratman (1999) and Kutz (2000), shared intentions can be understood in terms of “ordinary” individual intentions and does not require positing metaphysically mysterious group entities who hold the shared intention. Those on the other side of the debate like Gilbert (2006, 2009), Velleman (1997), and Meijers (2003) disagree on the grounds that individual or personal intentions cannot explain the obligations we seem to incur when we are engaged in joint action. According to these views, making sense of these obligations requires that we accept that shared intentions are basic and cannot be reduced to individual intentions.

I then turned to Baier (1997) and Stoutland’s (1997) arguments that there is a general individualist bias in the joint action literature, even among accounts that reject reduction. For

Baier and Stoutland, this is because most accounts assume that individual action is simple, or basic, and that joint action can be explained in terms of individual action but not vice versa. To that end, the accounts begin with individual action and try to determine what must be added to individual action to create joint action, or they assume that joint action is ultimately reducible to individual intentions. Standard accounts in the joint action literature also focus their analyses on cases that feature neutral act-types like painting houses or taking walks that are more compatible with these individualist assumptions. However, as Baier and Stoutland point out, this narrow focus risks painting an incomplete picture of the phenomenon, one that suggests that joint action only takes place between cognitively sophisticated agents who work cooperatively towards a shared goal.

More recently, several philosophers (Tollefsen 2005; Butterfill 2012; Pacherie 2013; Story 2021) have criticized the standard views on the grounds that they exclude young children. This is because leading accounts of joint action generally require participants to form complex intentions that are responsive to the mental states of their co-participants, an ability which presupposes a robust theory of mind. Young children (under four-to-five years old) lack a robust theory of mind and cannot form these complex intentions; however, they do appear to engage in joint action. In response, these recent accounts of joint action in children aim to formulate conditions for engaging in joint action that are less cognitively demanding but nevertheless explain how it is that children engage in joint action. This recent interest in how joint action in children might deviate from the standard structure is important because it paves the way for understanding some of the other neglected forms of joint action. For example, the minimal conditions used to explain joint action in children might be used to explain spontaneous forms of joint action in adults, which lack the pre-planning that seems built into standard accounts.



My dissertation proceeds in the same general spirit as the recent accounts of joint action in children. It aims to contribute to the literature on joint action not by attempting to tackle metaphysical issues about the nature of the intentions that seem to set joint action apart from other kinds of action, but instead by analyzing some atypical cases of joint action found in speech and early infant development. In this way, I contribute to a more complete understanding of the phenomenon, one that does not exclude actions that involve competition, antagonism, or a lack of shared interests among participants.

## References

- Baier, Annette. 1997. "Doing things with Others: The Mental Commons." In *Commonality and Particularity in Ethics*, edited by L. Alanen, S. Heinämaa, and T. Wallgren, 15-44. New York: St. Martin's Press, Inc.
- Bratman, Michael. 1992. "Shared Cooperative Activity." *The Philosophical Review* 101:327–41.
- Bratman, Michael. 1993. "Shared Intention," *Ethics* 104 (1): 97–113.
- Bratman, Michael. 1999. *Faces of Intention*. Cambridge: Cambridge University Press.
- Butterfill, Stephen. 2012. "Joint Action and Development." *The Philosophical Quarterly* 62 (246): 23-47.
- Chant, Sara Rachel. 2007. "Unintentional Collective Action," *Philosophical Explorations* 10 (3): 245–56.
- Clark, Herbert H. 2009. *Using Language*. Cambridge University Press.
- Davidson, Donald. 1980. *Essays on Actions and Events*. Oxford: Oxford University Press.
- Gilbert, Margaret. 1990. "Walking Together: A Paradigmatic Social Phenomenon." *Midwest Studies In Philosophy* 15 (1): 1–14.
- Gilbert, Margaret. 1996. *Living Together*. Lanham, MD: Rowman & Littlefield.
- Gilbert, Margaret. 2000. *Sociality and Responsibility*. Lanham, MD: Rowman & Littlefield.
- Gilbert, Margaret. 2006. *A Theory of Political Obligation: Membership, Commitment, and the Bonds of Society*. Oxford: Oxford University Press.
- Gilbert, Margaret. 2009. "Shared Intention and Personal Intention," *Philosophical Studies* 144:167–87.
- Grice, H. P. 1957. "Meaning." *The Philosophical Review* 66 (3): 377–88.

- Hammond, Paul. 2016. "Distinguishing Joint Actions from Collective Actions." *Synthese* 193 (9): 2707-20.
- Hume, David. 1748 (2007). *An Enquiry Concerning Human Understanding*. Cambridge: Cambridge University Press.
- Kopec, Matthew, and Seumas Miller. 2018. "Shared Intention is not Joint Commitment." *Journal of Ethics and Social Philosophy* 13 (2): 179-89.
- Kutz, Christopher. 2000. "Acting Together." *Philosophy and Phenomenological Research* 61 (1): 1-31.
- McMahon, Christopher. 2001. *Collective Rationality and Collective Reasoning*, Cambridge: Cambridge University Press.
- Meijers, Anthonie. 2003. "Can Collective Intentionality Be Individualised?" In *John Searle's Ideas about Social Reality: Extensions, Criticisms and Reconstructions*, edited by D. R. Koepsell and L. S. Moss, 168-83. Oxford: Blackwell.
- Miller, Seumas. 2001. *Social Action: A Teleological Account*. New York: Cambridge University Press.
- Moran, Richard. 2018. *The Exchange of Words: Speech, Testimony, and Intersubjectivity*. Oxford: Oxford University Press.
- Moll, Henrike, and Michael Tomasello. 2007. "How 14- and 18-Month-Olds Know What Others Have Experienced." *Developmental Psychology* 43 (2): 309-17.
- Pacherie, Elisabeth. 2013. "Intentional Joint Agency: Shared Intention Lite." *Synthese* 190 (10): 1817-39.
- Reid, Thomas. 1788 (2010). *Essays on the Active Powers of Man*. Edited by Knud Haakonssen and James A. Harris. Edinburgh: Edinburgh University Press.

- Searle, John. 1990. "Collective Intentions and Actions." In *Intentions in Communication*, edited by P. Cohen, J. Morgan and M. Pollack, 401–15. Cambridge: MIT Press.
- Stalnaker, Robert 2002. "Common ground." *Linguistics and Philosophy* 25 (5-6): 701-21.
- Story, Daniel. 2021. "Joint Action Without Robust Theory of Mind." *Synthese* 198 (6): 5009-26.
- Stoutland, Frederick, 1997, "Why are Philosophers of Action so Anti-Social?" In *Commonality and Particularity in Ethics*, edited by L. Alanen, S. Heinämaa, and T. Wallgren, 45-74. New York: St. Martin's Press.
- Tollefsen, Deborah. 2005. "Let's Pretend! Children and Joint Action." *Philosophy of the Social Sciences* 35 (1): 75-97.
- Tugendhat, Ernst. 1982. "The Employment Rule of an Assertoric Sentence." In *Traditional and Analytical Philosophy: Lectures on the Philosophy of Language*, translated by P.A. Gorner, 186-200. Cambridge: Cambridge University Press.
- Tuomela, Raimo. 2005. "We-Intentions Revisited." *Philosophical Studies* 125 (3): 327–69.
- Tuomela, Raimo, and Kaarlo Miller. 1988. "We-Intentions." *Philosophical Studies* 53 (3): 367–89.
- Velleman, J. David. 1997. "How To Share An Intention." *Philosophy and Phenomenological Research* 57 (1): 29–50.
- Wellman, Henry M. 2002. "Understanding the Psychological World: Developing a Theory of Mind." In *Blackwell Handbook of Childhood Cognitive Development*, edited by Usha Goswami, 167-87. New Jersey: Blackwell Publishing.

## Chapter 2: Telling as Joint Action

### 0. Introduction

In addition to the ordinary evidential way of coming to know things—for example, we may consult a weather forecast or check how the pedestrians outside are dressed to gauge the weather—we also come to know much of what we know second-hand, or on the basis of what we are told by others. Seeing that you have just come in from outside, I may simply ask whether it is still chilly outside and come to know that it is not when you tell me that it has warmed up. Some philosophers believe that acts of telling are the primary way in which we make what we know available to others, and for this reason, telling is often investigated as the main vehicle of testimonial knowledge. In this chapter I argue that the speech act of telling, that is, the means by which we transmit testimonial knowledge, is a form of joint action.

The idea that telling is a form of joint action is in one way unsurprising, because acts of telling, like paradigm forms of joint action, seem to require some degree of cooperation.<sup>13</sup> When you tell me and I come to know that it has warmed up outside, we both coordinate in various ways to make this possible. For example, you speak in a language I understand and pitch your voice at a level that is perceptible to me. In turn, I attend to your words and try to determine the content and force of your utterance. According to J. L. Austin, the role that I play as an interlocutor is crucial because telling is an “illocutionary” act, or an act in which speakers aim to get their interlocutors to recognize their intended meaning. As Austin observes, this seems to make the interlocutor’s recognition of the speaker’s act an essential part of the act itself, and so we may doubt that an act of telling occurs if it is not heard or understood by anyone (1962, 22).

---

<sup>13</sup> See Gelfert (2018) for discussion of the idea that testimony requires epistemic cooperation.

Some philosophers and linguists have also noted that cooperative practices in linguistic contexts are not unique to these contexts but are simply instances of what goes on in any rational cooperative endeavor. Robert Stalnaker (2014) writes that common ground, or the information that participants in a communicative exchange share and know they share, “play[s] a guiding role in any cooperative project, whether [it involves] language or not” (6).<sup>14</sup> Similarly, Mitchell Green (2017) observes that “even outside of language,” for example, when we adjust our paths to avoid colliding with one another, “we can accommodate each other in service to cooperative ends” (1589). Michael Bratman (1999) also notices the parallels between linguistic and non-linguistic contexts when discussing his conditions for shared cooperative activity, noting that they are “to some extent in the spirit of Grice’s classic discussion of [speaker] meaning” (101, n.13). Finally, Herbert Clark (1996) defends the stronger claim that “many phenomena [that] have been treated as features of language use” are really features of joint action, including “coordination, cooperation, conventions, turns, closure, joint projects, opportunistic actions, and the accumulation of common ground” (388).

These considerations—that acts of telling are to some degree cooperative and that cooperative principles in linguistic contexts are simply instances of broader cooperative practices—seem to pave a clear path to the idea that telling is related to the cooperative activities discussed in the joint action literature. However, I argue that it is a mistake to straightforwardly apply standard analyses of joint action to the case of telling. For example, in his recent work on the transmission of knowledge, John Greco (2020) rightly highlights the relationship between telling and joint action or joint agency but overlooks the problems that arise when analyzing

---

<sup>14</sup> Notably, Stalnaker claims to be inspired here by H.P. Grice (1957), who tried to give an account of speaker meaning that was independent of any conventional rules of language.

typical examples in the joint action literature—e.g., taking walks or painting houses together. As I will argue, telling, unlike the activities discussed in the joint action literature, can (1) occur in antagonistic contexts and so (2) need not involve cooperation around a certain kind of shared goal.<sup>15</sup>

Some of the building blocks of my argument that telling is a form of joint action are drawn from Richard Moran's (2018) recent book, *The Exchange of Words*. In the book, Moran is primarily interested in the question of how an act of telling can provide a particular kind of reason for belief.<sup>16</sup> He argues that this process is irreducibly social and intersubjective because "the status of [an] utterance as a reason to believe [...] depends on the speaker and her audience being related in a way with a particular structure" (xi). I believe that this structure is key to seeing why telling is a form of joint action, however, Moran stops short of calling the act of telling a joint action, in part because he believes that joint actions are cooperative in a way that acts of telling need not be. He observes that acts of telling occur even "[w]hen one person is threatening or verbally abusing another" but that it seems "misleading to call this 'joint action'" (2021, 723-24).

My goal in this chapter is to show that telling is a form of joint action, but that Moran's concerns reflect a need for an analysis of joint action that does not presuppose a certain kind of cooperation. I begin by considering Greco's (2020) argument that the transmission of knowledge

---

<sup>15</sup> Here, I don't want to deny that participants in a communicative exchange cooperate around a shared goal. However, these goals will sometimes be very general and appear only at the structural level of the exchange. For example, participants may only share the goal of communicating, or, in even more antagonistic cases, of speaking in the same language. But this sort of cooperation is different from the sort of cooperation that exists between two people who share a common interest in a particular topic of conversation. A similar point is made by Thomason (1990, 355-56) in his discussion of the difference between domain (content or subject matter) goals and discourse (structural, or conversational direction) goals.

<sup>16</sup> That is, Moran focuses on the transmission of reasons for belief from one person to another, rather than the transmission of knowledge.

is a form of joint agency as well as a similar proposal made by Krista Lawlor (2021). I argue that neither account successfully captures why the speech act of telling is a form of joint action. Instead, I argue that what is needed is the work of Ernst Tugendhat (1982) in addition to Moran's insight that telling essentially involves a speaker and interlocutor attending to the very same proposition that the interlocutor can respond to with affirmation, denial, or abstention. When we consider this fact, we see that in an act of telling, the speaker and interlocutor are both engaged in negotiating and determining what gets added to the common ground of the exchange.

I conclude by discussing some of consequences of the idea that telling someone something is a form of joint action. While such an account is not necessarily in tension with prominent analyses of shared intention, it does effectively broaden the category and our understanding of joint action by providing an analysis that does not build a particular kind of cooperation into joint action. This makes room for understanding how it is that interactions that are primarily competitive or even antagonistic might nevertheless constitute joint action.

### **1. Greco's argument**

According to Greco (2020), the transmission of testimonial knowledge is a form of joint agency. He offers a two-part defense of this claim: First, like others working in the epistemology of testimony, Greco suggests that the speech act of telling "is tailor-made for sharing knowledge that *p*" (48). Second, he argues that sharing knowledge that *p* involves joint action, because "the characteristic intention of the relevant speech act of telling – that of sharing knowledge – is also the shared intention of the relevant joint action" (49). I turn to shared intentions and why they are typically thought to be central to analyses of joint action in the following section. In this section, I summarize the first part of his defense.



Greco sketches an account of telling as “letting know” that is drawn from recent work on the relationship between telling and testimony by Elizabeth Fricker (2012) and Richard Moran (2018). From Fricker, Greco adopts the following four principles: First, in felicitous cases of telling, the speaker rightly takes herself to know that *p* and intends to share her knowledge of *p* with her interlocutor. That is, a speaker who does not take herself to know that *p* or is otherwise unintentionally uttering that *p* is not ordinarily taken to have felicitously told her interlocutor that *p*. Second, the speaker can accurately be described as trying to “let someone know” that *p* when she is engaged in the act of telling. Third, there is appropriate *uptake* when the interlocutor accurately identifies the content of force of the speech act and *consummation* when the interlocutor forms a belief on the basis of being told. And, fourth, telling is a social institution that is responsible for the transmission of knowledge (2020, 50).

From Moran, Greco adopts the idea that completed acts of telling (acts of telling where there is uptake) involve a kind of intersubjective dependence and mutual understanding between the speaker and interlocutor (2020, 51). According to Greco, the intersubjective dependence is due to the fact that, like other social acts, telling requires the input of another person (presumably in addition to the speaker) who recognizes the speech act and thereby plays a role in its completion. Telling also involves shared or common knowledge between the speaker and interlocutor because the speaker’s practical knowledge of what she is doing depends on her interlocutor recognising her as being engaged in the act of telling. That is, because the interlocutor’s recognition plays a role in the completion of the act of telling, the speaker’s

practical knowledge of what she is doing when she tries to tell him something must incorporate the role that the interlocutor plays in that act.<sup>17</sup>

The second part of Greco's argument is that the transmission of knowledge, which takes place through the speech act of telling, is a form of joint agency between a speaker and hearer who share the intention of "sharing knowledge," or "letting know," or "giving to know" (48). In order to understand this claim, we first need to understand what it is for two or more people to exercise their agency jointly. To do so, I will present Michael Bratman's (1999) account of shared intention as representative of the standard view, both because it is one of the most influential accounts in the literature and because Greco's account is largely drawn from Bratman's. Later, I will argue that because Greco uses a Bratman-style account of joint agency, he fails to capture why we should think of telling as a form of joint action.

### **1.1 Joint agency and Bratman's account**

Suppose you and I regularly meet at a particular time in order to take the bus to campus together. We board the bus together, decide on a pair of adjacent seats, and sit together for the duration of the ride. Many will accept that this is a paradigmatic form of joint action, but what is it that makes our action *joint*? It cannot be a difference in our *external* actions—our sitting together and riding the bus until we get to our shared destination—because these actions are shared with strangers on the bus who are also travelling to campus and occupying adjacent seats. It might not even be signaled by our willingness to engage one another in conversation, because

---

<sup>17</sup> Moran summarises the point like this: "For the speaker to be able to say, "I promise," "I claim," etc., her reflective understanding of her action must incorporate the point of view of her interlocutor and assume that he is indeed playing his part for the completion of the speech act [...] The speaker's knowledge of what she is doing in speech depends on this very knowledge being taken up and shared by her interlocutor; otherwise she will not in fact be doing what she announces in saying, I apologize, accept, refuse, tell" (2018, 34).

this condition seems neither necessary nor sufficient for our joint endeavor. It is not necessary because we might regularly decide to use the time to catch up on our individual work and therefore pass the time silently. It is not sufficient because I may strike up conversations with strangers on the bus without thereby being engaged in the joint activity of taking the bus with those strangers. Still, it seems plausible that, even in those instances where there is no outward difference between our actions when we take the bus together and the actions of the strangers on the bus, we take the bus together in a way that we do not with the strangers on the bus.

Because there is not necessarily an external difference between cases in which agents act together and cases in which their actions happen to coincide, the distinguishing feature of joint action is plausibly thought to be an internal one (Searle 1990, 402; Roth 2017). Specifically, many commentators in the literature believe that we can distinguish between joint and merely coincident action by looking to participant intentions.<sup>18</sup> This is suggested by the fact that if someone asks what I am doing as I wait with you at the bus stop, I am likely to say, “We’re taking the bus to campus,” which seems to express an intention to take the bus together.<sup>19</sup> However, if strangers who are also waiting at the stop were asked what they are doing, they are more likely to express intentions that describe individual actions.

These preliminary observations also suggest that the characteristic intentions of joint action are in some sense *shared* among participants. When I am asked about what I am doing, I describe an action that we intend to do together. I do not say: “I am going to campus, and she is, too,” even though this would not be an inaccurate description of our external actions. Instead, I say that *we* are going to campus and thereby indicate that the expressed intention is to be

---

<sup>18</sup> E.g., Alonso (2009); Gilbert (1990); Searle (1990); Bratman (1999); McMahon (2001); Miller (2001).

<sup>19</sup> Or alternatively; “So-and-so and I are taking the bus to campus.”

attributed to us both. But what is it to share an intention in a way that is characteristic of joint activity?

One prominent response to this question comes from Bratman (1999), who grounds his approach to shared intention in his general planning theory of intention. For Bratman, intentions are “elements of stable, partial plans of action concerning present and future conduct” that we make use of as rational agents to help realise our goals (2). Like individual intentions, shared intentions are elements of plans of action that serve to organize and coordinate our conduct around a goal. Bratman identifies three functions that shared intentions serve in the context of our shared activity.

The first function of shared intentions is to coordinate our activity with the activity of those with whom we act. If you and I intend to take the bus to campus together, we must settle on a time, a particular bus and line, and perhaps seats. The second function is to coordinate our planning about our particular roles in the joint activity. For example, if one of us occupies a window seat with easier access to the bell cord, that person should probably be the one who uses it to signal to the driver that we would like to stop, while the occupant of the aisle seat should begin to gather their things so we can exit the bus in a timely way. Finally, our shared intention to take the bus together will structure whatever bargaining we may need to engage in to resolve any conflicting preferences about, for example, the time at which we take the bus.

For Bratman, these functions are served by a set of interrelated individual intentions. Specifically, he believes that we share an intention to perform some action 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 1(a), 1(b), and meshing subplans of 1(a) and 1(b); you intend that we J in accordance with and because of 1(a), 1(b), and meshing subplans of 1(a) and 1(b).
3. 1 and 2 are common knowledge between us. (Bratman 1999, 121)

In other words, on this view, you and I share an intention to take the bus together if we both intend to take the bus together, are both aware of the other's intention that we take the bus together, and indeed intend to take the bus together in part because of the other's intention that we take the bus together. Further, we should both intend to mesh our subplans for taking the bus together (about the time we take the bus, the line we take, the stop we walk to, etc.) and negotiate to make these subplans consistent if they are not. Bratman believes that participants will do so in part because of rational pressure to make their goal, e.g., taking the bus together, means-end coherent (125). Thus, on this account, agents form interlocking intentions and rationally cooperate in various ways to jointly realise a goal that they happen to share.

In the next section, I turn to how Greco uses this general view to defend his claim that telling is a form of joint agency in which speakers and interlocutors share an intention of "sharing knowledge," "letting know," or "giving to know."

## **1.2 Shared intentions and the transmission of knowledge**

Greco's view of joint agency is based on four features about which he believes there is broad agreement in the joint action literature. For Greco, joint action (1) involves shared intentions, (2) involves sub-plans, (3) is interactive in the sense that each participant is responsive in various ways to the actions of the other, and (4) is interdependent in the sense that the action cannot be carried out alone but depends on the contributions of each participant (2020,

56). Thus, Greco straightforwardly adopts features (1) and (2) from Bratman, while (3) and (4) reflect Bratman's claims about the interdependence of participant intentions as well as the rational pressure to mesh subplans. From these features, Greco develops the following view of joint agency operative in knowledge transmission:

**(KT)** Knowledge that *p* is transmitted from a speaker S to a hearer H just in case S successfully tells H that *p*.

And *that* happens just in case: (1) S knows that *p*; (2) S asserts that *p* with the intention of sharing knowledge that *p* with H; (3) H understands and shares S's intention; (4) S and H act jointly so as to bring about their shared intention (i.e., so as to consummate the speech act in condition 2). (Greco 2020, 57; Greco's emphasis)

Before assessing this argument, it is worth noting that Greco locates the joint action in *successful* (or "consummated," in Fricker's terminology) transmissions of knowledge, rather than in a completed act of telling. Thus, we have a case of joint agency in Greco's view when the interlocutor believes what the speaker says and thereby comes to know what the speaker in fact knows. This is because Greco is primarily concerned with the transmission of knowledge, something which can only take place in successful acts of telling, where the speaker knows that *p*, tells the interlocutor that *p* and the interlocutor thereby comes to know that *p*. My view, however, is that completed acts of telling—acts of telling that are recognized by an interlocutor—constitute joint action. This means that acts of telling do not need to be successful to constitute joint action: the speaker may not in fact know that *p* (she may be mistaken about whether *p* is true) or the interlocutor may recognize the speaker's speech as an act of telling but fail to believe what she says and therefore fail to come to know what she knows. Since Greco

presumably acknowledges that unsuccessful but completed acts of telling take place, his target is properly a subset of acts of telling.<sup>20</sup> However, even if we grant that Greco is concerned with a separate or additional joint action that takes place only when the speaker is believed and the interlocutor comes to know what she knows, his account faces serious problems. In what follows, I raise some worries about Greco's account as well as a similar proposal made by Krista Lawlor (2021).

## 2. Telling and the standard view of joint action

As we saw in the last section, Greco argues that successful acts of telling involve joint agency because the speaker and interlocutor must share the intention of “sharing knowledge” or “letting know” to facilitate the transmission of knowledge. If Greco is right, his account would have the advantage of easily slotting into the joint action paradigm. However, I believe that there are issues with analyzing telling in the way that actions like taking a bus together or painting a house together are analyzed in the joint action literature. These issues are evident in condition (3) of KT—that H understands and shares S's intention.

First, it is not clear why we should think that the *interlocutor* shares the intention of sharing knowledge with the speaker, even in those cases where the interlocutor does in fact believe what the speaker says. That is, we might agree that when a speaker tells her interlocutor something, she intends to share her knowledge of or let him know that *p* but doubt Greco's claim that the interlocutor shares the very same intention. The interlocutor, we might think, can be

---

<sup>20</sup> Greco's claim is therefore slightly different from mine because he is trying to defend the view that the transmission of knowledge involves shared agency. However, because the transmission of knowledge is essentially connected to telling, he thinks that “the speech act of telling and the joint action of sharing knowledge are importantly related” (Greco 2020, 49). The difference between our views is that mine applies to all acts of telling, not just the successful ones where knowledge transmission in fact occurs.

plausibly thought of as having the intention of listening to the speaker, or learning something new, but not the intention of sharing knowledge.

Greco anticipates this sort of objection and responds by arguing that “telling,” as well as acts like “promising,” “betting,” and “passing a ball,” have both individual and joint senses (2020, 60). That is, these terms sometimes refer to the actions of individuals, but can also refer to joint actions. Greco illustrates the point with the following example:

The distinction is easy to see in the case of passing a ball. On the one hand, there is Brady’s individual action of throwing the ball, which he could do himself by throwing into a practice net. On the other hand, there is “passing the ball” in the sense of completing a pass, something that Brady and Gronkowski do together. In this sense, if Gronkowski does not catch the ball, then Brady was trying to pass the ball to Gronkowski but didn’t. (Greco 2020, 60)

The idea here is that if we object that the interlocutor does not share an intention with the speaker in an act of telling, it is because we are thinking of telling in its individual sense, where it is something that is done by the speaker. However, there is also a joint action sense of telling in which the act does not occur unless it is completed by an interlocutor. That sense of telling, for Greco, involves shared intention.

But even if we agree that there is a social sense of terms like telling or passing a ball—and it seems likely that these terms are primarily social<sup>21</sup>—we may nevertheless object to

---

<sup>21</sup> Greco’s distinction between individual and social senses of telling seems inaccurate. Actions like passing the ball or betting do not in fact appear to have the individual sense that Greco claims they do. Instead, they appear to belong to Thomas Reid’s category of primarily social acts like marrying, buying, contracting, or telling, which require at least two people to take place. In the case of passing a ball for example, we do not say that someone *passes a ball* if they are alone and happen to be tossing a ball around. This is evident in the way that Greco himself switches to ‘throwing the ball’ when trying to capture the individual sense. The term passing the ball, on the other hand, seems to necessarily introduce another person to whom the ball is passed. Similarly, speakers do not



Greco's claim that the speaker and interlocutor in an act of telling share an intention. That is, we might agree that Brady does not pass the ball to Gronkowski if Gronkowski does not catch it, but not accept that Gronkowski shares the intention of *passing the ball* with Brady in the same way that you and I share the intention of *taking the bus together*. In the case of taking the bus, you and I both intend to take the bus together because we are each going to take the bus. However, in the case of passing the ball, Brady and Gronkowski seem to intend different things, because they are engaged in different actions. While they are both needed to complete the action of passing the ball, Gronkowski's role in the action of passing the ball is to catch it, thus, he is more naturally described as intending to catch the ball or complete the pass.

Still, we might think that the difference between the cases of passing the ball or telling and taking the bus together is merely grammatical. Passing the ball and telling are forms of joint action that do not take the form of "we are *x-ing together*," but this does not mean that Bratman-style shared intentions do not apply. Greco does not specify how it is that the hearer H in his third condition "understands and shares S's intention," but perhaps there is a plausible story to be told about why this is only superficially different from how we share an intention when we take the bus together. However, even if we grant this point, there is a further issue with condition (3).

The issue stems from the fact that Bratman's account of shared intention is fairly narrow in scope. By his own admission, it rules out cases involving coercion and does not explain fully explain cases that involve competition. For example, if we were to intend to take the bus together as a result of my threatening or otherwise coercing you into taking the bus with me, we would fail to meet Bratman's conditions for shared intention. This is because you would intend to take

---

complete acts of telling when they are alone. They may perform an assertion, or otherwise claim that something is the case, but they cannot engage in the act of telling without being recognised as doing so by an interlocutor.

the bus together because I threatened or compelled you into taking the bus together and not, as Bratman's second condition holds, because I intend to take the bus together. In such cases, Bratman believes that I do not treat you as a participating intentional agent because my forcing you to take the bus together bypasses your intentional agency. Competitive games also mark a limit of the explanatory power of Bratman's account, because although both participants will usually share an intention to play the game in question, they will not intend to mesh subplans, since this will undermine their individual goals of winning the game (1999, 122).

There are also further limits to Bratman's account. One has to do with the fact that sharing intentions in the appropriate way requires participants to formulate complex, interlocking intentions or intentions about intentions. Because it is generally agreed that understanding the intentions of others is a cognitively complex process, young children and people with certain intellectual disabilities will be unable to share intentions in the way that Bratman's account requires. That is, because sharing an intention requires not only that you and I both intend to perform some action together but also that we are each aware of this fact, and indeed intend to perform that action in part because of the other's intention, the process is too cognitively demanding for those who lack a robust theory of mind. And, because children only develop a robust theory of mind around four or five years old, and some people with intellectual disabilities have limited or impaired theory of mind, this would rule out the possibility that these groups share intentions in the way that Bratman describes.<sup>22</sup>

More generally, Bratman's account seems ill-equipped to explain any action that deviates from the broadly cooperative model in which two or more people share an intention to pursue a

---

<sup>22</sup> See Astington and Jenkins (1995) for discussion of the emergence of theory of mind in children, and Abbeduto et al. (2004) for discussion of the effects of intellectual disability on theory of mind.

shared goal. In Bratman's view, participants in a joint activity are not only concerned with acting with each other but also in ensuring that each participant's intention partly explains their performance of the joint activity (1999, 118). As we saw earlier, the case of coercion fails to meet this standard because the action is explained by the fact that one person threatens the other, and not just by the fact that each person intends to do the action. This suggests that the ideal case of shared intention is one in which two or more participants freely and equally contribute to the performance of the joint activity. If either participant attempts to steamroll the other and control the various subplans they jointly adopt, the shared intention may be dissolved (1999, 22).

This is not to say that Bratman's account fails to capture an important phenomenon. After all, his target is shared intention, which he only takes to be typical of a certain kind of unified or joint activity. This leaves open that some other kind of joint activity could take place in the absence of shared intention. Bratman seems to acknowledge this when he suggests that while a case in which I threaten you into intending that we go to New York together cannot involve a genuinely shared intention because it bypasses your intentional agency, such an action "will in a way be unified" though it will not be "the kind of unified agency [that is] characteristic of shared intention" (1999, 118).

The above considerations show that Bratman's account is really directed at explaining the cooperative pursuit of a shared goal between two people who are roughly equal and fairly cognitively sophisticated. Again, this is reflected in the way that coercion, which would introduce relations of power or authority and subordination, rules out the possibility of shared intentions. Similarly, the involvement of a child who does not yet have a theory of mind would also rule out the possibility of a shared intention, as would the involvement of a person with profound intellectual disabilities. Yet, intuitively, children and people with intellectual

disabilities do engage in joint action. Therefore, it seems fair to say that Bratman's account excludes a wide range of possible cases of joint action, and there is no reason to think that plausible cases of joint action must fit into this shared intention framework.

So, it is a mistake to think that joint action necessarily involves shared intention, but why think that telling, in that case, is a form of joint action that does not fit into the shared intention framework? The main reason is that telling is compatible with the situations described above. Speakers can tell their interlocutors something and be believed even in those cases that Bratman explicitly rejects because they involve coercion or other forms of antagonism. Similarly, children both tell and are told things long before they have the sort of robust theory of mind required to form complex interlocking intentions with other people. This suggests that Bratman's account is in tension with the claim that telling involves joint agency, because acts of telling seem to occur in contexts that shared intentions do not.

A similar problem arises in Krista Lawlor's brief proposal for understanding telling as a form of joint action in a review of Moran's *The Exchange of Words*, because while she does not draw on standard shared intentions, her proposal seems assume a similar kind of cooperation. Borrowing from Herbert Clark (1996), Lawlor suggests that telling involves a speaker and interlocutor who jointly propose to change the interlocutor's mind:

“As (Clark 1996) spells out the idea, speech acts involve a proposal for joint action by the speaker, taken up or modified by the addressee; in the case of telling, the speaker S proposes to the addressee A that they jointly change A's beliefs about p, and if the audience accepts the proposal, he gets a reason to believe P, but if he rejects the proposal, no reason to believe P is constituted by

the speaker's utterance. (Moran may prefer to say that S proposes that they jointly provide A with a new reason to believe p.)" (Lawlor 2021, 703)

But it is strange to think of interlocutors as necessarily accepting a proposal to jointly change their minds in every act of telling. In some cases, particularly instances of telling that occur during arguments or other forms of antagonism, interlocutors will be entirely uninterested in changing their minds, especially *before* having and considering the reason for belief offered by the speaker in the way Lawlor suggests. Further, many acts of telling occur spontaneously and between speakers and interlocutors who lack, or do not know if they have, overlapping goals. Consider, for example, the case of catcalling, or any exchange into which an interlocutor might be unexpectedly drawn. In those cases, the interlocutor may want nothing more than to end the exchange, and yet, successful acts of telling may nevertheless occur.

So far, I have been arguing that Greco's third condition fails because it draws on a notion of shared intention that seems ill-equipped to explain activity that is not primarily cooperative. But it is important to remember that the above considerations are not aimed at showing that speakers and interlocutors do not cooperate at all when engaged in more fraught cases of telling. Even in these cases, speakers and interlocutors will often both intend to communicate with one another (even if one person merely wants to communicate that they wish to no longer communicate) and will do things to facilitate this like speak in a language and at a pitch that is intelligible to their counterpart in the exchange. Thus, it is plausible that speakers and interlocutors do cooperate on a more structural level in these interactions. However, these intentions are secondary, background intentions that do not capture a distinctive shared intention of the act of telling.

In sum, while Greco fruitfully highlights some of the ways in which acts of telling depend on the contributions of both the speaker and interlocutor, his argument that this process necessarily involves Bratman-style shared intentions is mistaken. As I have shown, Bratman's account of shared intention is aimed at explaining actions in which participants cooperate around a shared goal that is closely related to the task at hand. For example, in addition to cooperating around the more structural goal of speaking English, two agents who share intentions in Bratman's view would cooperate around the goal of discussing the weather. This would rule out the possibility of shared intentions in cases involving coercion or competition, but intuitively, acts of telling are compatible with these situations. Bratman-style shared intentions also seem to require that participants understand and be responsive to the intentions of others, which in turn seems to require a robust theory of mind. However, even children and others who lack a robust theory of mind seem to engage in acts of telling. For these reasons, telling does not seem to involve a shared intention of "sharing knowledge" or "letting know," or "giving to know," and so on. Thus, if telling is indeed a form of joint action, we need a different account of why that is.

### **3. Telling and shared understanding**

In the previous section, I discussed and pointed out some issues with the strategy of analyzing telling in the same way that the primarily cooperative examples are analyzed in the joint action literature. Because acts of telling regularly occur in situations that involve coercion and antagonism, they cannot be explained by the shared intentions that are characteristic of primarily cooperative examples, because these intentions are incompatible with coercive or antagonistic contexts. However, I believe that there is still a case to be made in defense of the claim that telling is a form of joint action.

In what follows, I present my positive view according to which telling is understood as the speaker and interlocutor's joint constitution of their common ground. To do so, I draw on Moran's assurance view of telling, Ernst Tugendhat's argument against the unilateral understanding of acts of assertion, and Robert Stalnaker's notion of common ground. With these tools, I advance a view of telling as a form of joint action that does not rely on standard analyses of shared intention in the joint action literature.

### 3.1 Moran's assurance view of telling

Moran's (2018) account of telling includes several of the features already discussed. For example, he believes that telling is the paradigmatic way in which we transmit knowledge to others and that acts of telling are *completed* when they are recognized by the interlocutor and *successful* when the interlocutor believes what he is told. He also believes that acts of telling involve a speaker who takes herself to know that  $p$  and tells her interlocutor that  $p$  with the intention of informing the interlocutor that  $p$ . Unlike Greco, however, Moran does not believe that the transmission of knowledge is facilitated by the speaker and interlocutor engaging in a form of joint agency. Instead, for Moran, speakers facilitate this process by offering their interlocutors an *assurance* that  $p$  is true.

Assurances are provided when a speaker invites the interlocutor to rely on her for the truth of  $p$  by "explicitly [assuming] a certain responsibility" for her interlocutor's belief that  $p$  if he does indeed believe her (44). In turn, the interlocutor gains certain rights of complaint when he accepts the speaker's assurance that  $p$  but  $p$  turns out not to be true.

Of course, quite a lot of the things we say do not seem bound up in these ideas of assurance or responsibility-taking. Many of these utterances will be assertions—the broad

category of speech-acts in which a speaker claims that something is the case. Because we can claim that something is the case for a myriad of reasons, many acts of assertion will not be attempts to transmit our knowledge of some  $p$  to an interlocutor. A simple case of this occurs when a student recites some information back to a teacher. In such cases, the speaker does not necessarily take herself to know or believe that  $p$ , nor does she believe that she is informing her teacher that  $p$ .

When a speaker does tell an interlocutor that  $p$  with the intention of informing him that  $p$ , however, she offers him an assurance that what she says is true. Many of the details about how this occurs are beyond the scope of this chapter and not necessary for understanding Moran's basic argument. However, it is important for our purposes that Moran takes these assurances to provide a particular kind of reason for belief that he calls a *testimonial* reason for belief. For Moran, testimonial reasons for belief are provided when the speaker and interlocutor are related to one another in a particular way—one that allows them to achieve a shared understanding of the proposition under consideration.

In the next two sections, I describe the nature of testimonial reasons and how they give rise to a particular kind of shared understanding. The shared understanding that obtains between the speaker and interlocutor is a key part of my argument that telling should be understood as a form of joint action even though it is incompatible with standard shared intentions.

### **3.2 Testimonial vs. evidential reasons for belief**

In an ordinary situation in which a speaker tells her interlocutor something, her words provide him with a particular kind of reason for belief. For Moran, this reason for belief is distinct in important ways from the kind the interlocutor might gain when he attends to her words



in the way a detective might attend to a handkerchief that was dropped at the scene of a crime—as a state of affairs with evidential value.<sup>23</sup> When an interlocutor seeks this evidential kind of reason from a speaker’s act of telling him that “*p*,” he uses her act of telling—like the detective uses evidence left at a crime scene—to try to draw conclusions about the truth of *p*.

In cases like this, the fact that the speaker intentionally says that “*p*” might play a role in the reason for belief derived from the speaker’s words, but it is not *by itself* a reason for belief. The speaker’s intentional act might instead provide *evidence* that *p* is true, if the interlocutor also has further evidence: evidence, for example, that she reliably says things that turn out to be true. But, as evidence, her act of telling is not something that by itself provides a reason to believe *p*. That is, the interlocutor does not gain a reason to believe *p* from the speaker’s intentional act of telling him that “*p*.”

Moran thinks that something different occurs in the ordinary dialogic situation. Here, the speaker’s intentional act is sufficient to generate the reason for belief that is taken up by the interlocutor. And, in a successful act of telling, where the interlocutor believes what the speaker says, the interlocutor takes the speaker’s utterance as a direct reason for belief “because he believes *her*” (Preface, x). This suggests that testimonial reasons for belief involve believing some *p* that the speaker intentionally says because she tells him that *p* and not because the

---

<sup>23</sup> This example comes from Grice’s (1957) paper “Meaning.” In the paper, Grice sets out to articulate what constitutes non-natural speaker meaning, as opposed to what he calls natural meaning or meaning that attaches to causal or entailment relations that obtain between natural objects in the world (e.g., a particular kind of ‘spot’ on human skin might mean measles). For Grice, non-natural meaning is meaning that obtains when a speaker intends that her audience’s recognition of her intention that they come to believe some *x* play a necessary role in their coming to believe that *x*. Grice also hints that this condition is met when one person tells another something, and not simply when a speaker deliberately and openly lets her audience know that *x* is the case, with the aim of producing a belief that *x* in her audience. This suggests that Grice is alive to the distinction Moran is proposing: that telling someone something and the testimonial reason it involves is distinct from simply providing them with evidence, even if you make it clear to them that you’d like them to reach a particular conclusion. In the latter case, the audience may only derive an evidential kind of reason for belief, but Grice seems to say that this is not what is involved in telling someone something.

interlocutor knows that the speaker tends to say things that are true. Unlike the case in which the interlocutor deduces an evidential kind of reason from what the speaker says, telling someone something and being believed is marked by the interlocutor believing the very thing the speaker means to convey because he believes “*the speaker herself*” (Preface, x).

The distinction here can be illustrated with the following sort of case. Suppose that you are headed out for dinner with a few friends but there is some uncertainty about the location of the agreed upon restaurant. You happen to know where this restaurant is, and so you confidently say to them that the restaurant is two blocks north, on the west side of a particular street. Your friends hear you, and, importantly, do not *disbelieve* you, but they do continue to consult a map to locate the restaurant.<sup>24</sup> Here, I think it is not unusual to feel at least slightly affronted that your friends do not “take your word for it” and instead choose to continue to consult the map.<sup>25</sup> But why is that? If you are confident about the restaurant’s location, you know that your friends will soon independently verify what you have said. So, it seems it is not that you are concerned that your friends will not know that you have said something true, rather than false, because they will come to know that you were right when they check the map.<sup>26</sup>

There are, of course, a number of plausible explanations for your friends’ reaction to your testimony: your friends may know that you are hopeless at directions, or they may be reluctant to give up the kind of guarantee that Google Maps provides. Here, I only want to suggest that at

---

<sup>24</sup> If your friends merely disbelieve you, this could simply be an unsuccessful act of telling in which your friends recognise but reject the reason for belief you present to them.

<sup>25</sup> Cf. Anscombe “It is an insult and it may be an injury not to be believed” (1979, 151).

<sup>26</sup> Of course, there are a number of ways cases like this can go, and in some of them the speaker may well be most interested in being proven right. In a case like this she may smugly wait for the validation that she is certain will come when her friends consult the map. This is another kind of familiar exchange, but it does not rule out the kind of case I refer to here: cases in which we seem to want others to *believe us* directly.

least one of these plausible explanations is found in Moran's distinction between testimonial and evidential reasons. That is, the reason for belief that you try to offer is not taken up and your utterance is instead treated as mere evidence (in this case one piece among many) from which a reason for belief may then be deduced. Your attempts to provide a testimonial reason—or to *tell* your friends about the location of the restaurant—are therefore frustrated because your friends relegate your testimony to the status of evidence.

The testimonial reasons that are offered and taken up in acts of telling are important for our purposes because they give rise to a particular joint or mutual state between the speaker and interlocutor. Later I will argue that it is this state, rather than shared intention, that serves as the basis of the speaker and interlocutor's joint action in the act of telling.

### **3.3 Shared understanding**

For Moran, testimonial, but not evidential, reasons give rise to a shared understanding of the proposition that the speaker presents the interlocutor with a reason to believe. That is, a distinctive feature of the giving and receiving of testimonial reasons is that the speaker and the interlocutor will jointly consider the very same proposition. This allows for an intersubjective experience that evidential reasons do not provide. To see why, it is helpful to look at Ernst Tugendhat's (1982) argument that acts of assertion are not unilateral acts that depend on the speaker alone.

Tugendhat defends this argument by drawing on the following data: every act of assertion provides room for a linguistic response of the same kind. Consider the case in which you tell me that it is cold outside when you have just come in from outside. In such a case, it seems open to me to respond in a few different ways. If you tell me that it is cold outside after I ask you if it is

cold outside, I will likely accept your answer and believe that it is indeed cold outside. In other situations, however, I might instead deny your claim that it is cold outside. This might occur if I know we tolerate the cold differently, or because I did not agree with your past reports about how chilly it was outside. In yet other situations, I might respond with neither affirmation nor denial. In these cases, I may have some reasons to believe you and others not to, and so abstain until I have further reasons to tip the scale in one way or another. In sum, Tugendhat's point is that interlocutors can always respond with affirmation, denial, or abstention in response to an assertion.

This claim may seem strong to us because it seems unlikely that every act of assertion makes room for a response. However, I believe Tugendhat's idea can be made more plausible with two clarifications. First, it is unlikely that *every* act of assertion makes room for a linguistic response, for the reasons briefly discussed in the previous section. That is, because some acts of assertion do not seem to be informative, it is likely that not all acts of assertion make room for a linguistic response. An actor in a staged play may make several assertions while delivering a soliloquy, but these assertions seem not to make room for a linguistic response from the audience or the other actors. After all, the actor on stage is not trying to get the audience or the other actors to believe the propositional content of his assertions. However, the subset of acts of telling that are informative do seem to make room for linguistic responses from the interlocutor, who may respond with affirmation, denial, or abstention. The second clarification is that Tugendhat's argument is best understood as the claim that it is at least in principle possible for the interlocutor to respond. In certain situations, a linguistic response will not be practically possible, or interlocutors may prefer to avoid responding due to various norms of politeness.

The insight that it is open to the interlocutor to respond in kind to the speaker highlights the fact that the speaker and interlocutor jointly consider the very same proposition which the speaker puts forward and the interlocutor responds to with affirmation, denial, or abstention. This is usefully contrasted with what goes on in the case of evidential reasons, where the interlocutor unilaterally deduces a reason for belief from the speaker's words. In these cases, because the interlocutor derives his own reason for belief, he does not consider the very same proposition as the speaker. Instead of directly considering  $p$  because the speaker has provided a reason to believe  $p$  by assuring him that it is true, the interlocutor considers the state of affairs that consists in the speaker's claim that  $p$  is true. From there, the interlocutor may derive a reason to believe  $p$  if he has further evidence that the speaker tends to say true things, etc. However, because he does not accept the offer of a testimonial reason for belief, his consideration of  $p$  is not something that he does jointly with the speaker. Even if he eventually comes to believe  $p$ , it will be the result of a unilateral process in which he derives  $p$  from his evidence (which includes the state of affairs that consists in the speaker's claim that  $p$  is true) that  $p$  is true.

Speakers and interlocutors therefore share an understanding of the proposition in question in the giving and receiving of testimonial reasons. But how should we characterize this shared understanding or intersubjective experience that obtains in acts of telling? It is tempting to think that shared understanding can be captured by concepts like mutual or common knowledge that philosophers of language have used to characterize the information that is held in common between speakers and interlocutors in communicative acts. Common knowledge obtains between a speaker and interlocutor when they are both aware of some fact, are also each aware that the other knows this fact, and that the other knows that they know this fact, and so on. However, as Moran points out, common knowledge by itself does not seem to guarantee that the two minds

are jointly considering the very same fact. Instead, common knowledge appears consistent with the case above in which the interlocutor merely deduces a reason for belief from the speaker's words.

Consider a case in which Jack would like Sue to know that he has had a falling out with a family member. In such a case, we can imagine that Sue eventually comes to know that Jack would like her to know that he is upset about falling out with a family member because she has recognized his intention that she come to know about the falling out. We can also imagine that Jack knows that Sue knows this, and that Sue also knows that Jack knows that she knows, etc. The problem here is that it seems possible for Sue and Jack to come to have all of this information by unilaterally deducing what it is the other knows. Because Jack does not tell Sue that he is upset, she deduces an evidential reason to believe that he is upset from his actions or hints, but this does not place Sue and Jack in the kind of mutual encounter that goes on in acts of telling.

For Moran, this is reflected in the fact that there is no possibility of a linguistic response from Sue when she simply deduces that Jack wants her to know that he is upset. For Moran, this is because what Sue learns is "not something of the right logical type to be capable of truth or falsity itself" and therefore, not something that Sue could respond to with affirmation or denial (163). It would be like the Gricean detective responding to an evidential state of affairs like a handkerchief dropped at a crime scene by a suspect with affirmation or denial. This is because a state of affairs can only provide the interlocutor with an evidential kind of reason, but of course, this is the kind of reason we have established is distinct from the kind of reason we get when one person tells another something. Instead, making sense of the basic fact that the interlocutor

should be able to respond with an affirmation or denial requires that we conceive of the speaker as claiming that something is the case, something which might be true or false.

How then should we characterize the shared understanding that obtains between two people who are engaged in the giving and receiving of testimonial reasons? For Moran, in addition to being common knowledge between the speaker and interlocutor, the proposition in question must also be open or public between them. This is typically achieved when a speaker openly and explicitly tells her interlocutor that *p* in the paradigmatic face-to-face case of telling, though it may also be achieved in different forms as long as the requirement that the relevant proposition is open or public is met.

To see what Moran is getting at here, it is useful to briefly compare Gilbert's (1990) discussion about the relationship between common knowledge and acting together. In her paper "Walking Together: A Paradigmatic Social Phenomenon" Gilbert tries to determine what distinguishes two people who are taking a walk together from two people who happen to fall into step on the same path. She rejects the idea that what is distinctive about the case in which two people take a walk together is that they have common knowledge of the fact that they each have the goal of taking a walk together. That is, she rejects the idea that two people who are walking side by side are walking *together* just because they are aware that they each have the goal of walking together. She believes this is not sufficient to conclude that two people are indeed walking together in part because neither person would be entitled to complain if the other suddenly veered off in a different direction or decided to go home.<sup>27</sup> What would license this

---

<sup>27</sup> The point here can be obscured by the example since walking together might not be thought to involve much more than the action of walking side by side. But a similar point can be made about a case in which Jack and Sue are both aware (perhaps because mutual friends informed them, or they figured it out unilaterally) that they each have the goal of going on a date together. Even with common knowledge of this fact in place, it seems possible that Jack and Sue could e.g., get dinner together (or generally perform date-like activities) without (necessarily)

right of complaint? For Gilbert, it is when each person expresses their willingness and thereby commits themselves to forming what she calls a “plural subject” of the goal of taking a walk together, but the details of her argument here are not necessary for our purposes. The key point is that common knowledge of a goal to take a walk together is not what distinguishes a genuine case of walking together from a case in which two people simply happen to fall into step. Instead, like Moran, Gilbert believes that they must make their goal of walking together open or public between them, and in her view, this is also achieved via a mutual encounter in which they jointly commit themselves to pursue the goal of walking together.

In sum, the sense of shared understanding that obtains between the speaker and interlocutor in an act of telling is not, or not *merely*, the “common knowledge” that might obtain when the speaker knows that the interlocutor knows some  $p$ , and the interlocutor knows that the speaker knows that he knows that  $p$ , and so on. Instead, both Moran and Gilbert think that  $p$  must additionally be open or public between them. For Moran, this additional condition secures Tugendhat’s requirement that acts of telling involve the possibility of a response.

In the next section, I combine these insights from Moran and Tugendhat with the notion of common ground. With these ideas in place, we will finally be in a position to see why telling should be understood as a form of joint action.

#### **4. Telling as Joint Action**

When a speaker utters that  $p$  with the intention of informing an interlocutor that  $p$  is true and the interlocutor correctly recognizes her as being engaged in the act of telling, the

---

being on a date. Thus, common knowledge of a goal of performing some action  $x$  while behaving in a manner that is consistent with  $x$ -ing is not enough to count as a case of  $x$ -ing.



proposition  $p$  is unrestrictedly open between them in the sense that they mutually consider  $p$ . However, the mere fact that some proposition  $p$  is unrestrictedly open between the speaker and interlocutor does not amount to joint action. This does not yet show us that there is anything that the speaker and interlocutor together *do*, though it does establish Moran's thesis that telling someone something is intersubjective, because it involves both parties entertaining the very same proposition which is open between them. How then do we establish that telling is in fact a joint *action*? The final piece needed to answer this question is Robert Stalnaker's (2014) notion of common ground, the basic idea of which I sketch below.

For Stalnaker, common ground refers to a common store of information held by participants in a communicative exchange. More precisely, it is the information that participants mutually accept for the purposes of conversation (Stalnaker 2014).<sup>28</sup> These communicative exchanges are typically based on an *assumption* of common ground. For example, the person who comments "It's busy today, isn't it?" in a crowded grocery store assumes that her interlocutor also speaks English, is aware that they are both in a busy grocery store, and can at least work out from her comment that this is unusual, if her interlocutor is not already aware of that fact.<sup>29</sup> The interlocutor's ability to work out what the speaker says is also based on the interlocutor's assumption of common ground. The interlocutor assumes that the speaker speaks

---

<sup>28</sup> In turn, some information  $i$  is common ground between us just in case we both accept  $i$ , both accept that we accept  $i$ , both accept that we accept that we accept that we accept  $i$ , and so on. The iterative structure in Stalnaker's conception of common ground therefore has to do with what each party to a conversation accepts, where acceptance is the propositional attitude of *accepting* (that is, *treating as true*) for the purposes of conversation. While this structure is an important part of Stalnaker's theory, I try to give a common sense version of his main ideas due to limitations of space.

<sup>29</sup> There are many more assumptions of both a more general and specific nature and CG includes both linguistic and non-linguistic elements. For example, the speaker is unlikely to address her utterance to an infant because she assumes that her CG with the infant is limited, and the infant is unlikely to understand the significance of her utterance.

English and is not merely repeating a phrase that she does not understand, and that the speaker is referring to the store they are both in and the unusual number of people in it, etc. While the assumed common ground between speakers and interlocutors is unlikely to perfectly coincide, typically there is enough overlap to facilitate effective communication.

Different speakers and interlocutors have varying degrees of common ground between them, and this is often determined by their membership in various communities. Two people from the same culture can generally expect to have more common ground between them than two people from different cultures, and similarly, a family or friend group within a culture will have more common ground than groups of strangers from that same culture. But importantly, common ground is not static, and both linguistic and non-linguistic events can lead to changes in common ground. As a typical conversation between two or more people progresses and new information comes to light, the common ground will evolve in response to that information. Similarly, the common ground between a group of strangers waiting to cross a street will expand to include manifest events like the fender bender that occurs in the intersection as they wait.

Why think that common ground shows that telling is something that the speaker and interlocutor together do? A clue comes from Tugendhat's claim that telling always involves the possibility of a reply, because there is something that the interlocutor does *in* responding to the speaker. When the interlocutor responds with affirmation, denial, or abstention to the proposition offered by the speaker, he is also helping to determine whether this proposition will be added to the set of facts that will be presupposed in the evolving context of their exchange. Roughly, this means that the speaker and interlocutor are jointly engaged in the shared project of entertaining the fact the speaker attests to and deciding whether and how it is added to the common ground.

In practice, this means that when the interlocutor accepts the reason for belief offered by the speaker and believes that  $p$  is the case, the proposition in question is added to the common ground of the exchange. If the interlocutor recognizes and rejects or abstains from deciding whether to accept or reject the proposition put forward by the speaker, the common ground can also record that fact because, as Stalnaker puts it, common ground “includes both information about the subject matter of the discourse, and information about the discourse itself” (2014, 36). Thus, even when the propositional content of the act of telling is disputed and not added to the common ground, the fact of that disagreement is added, because both participants can accept this fact. If there is no uptake or recognition of the speaker’s utterance as an act of telling there will also be no change to the common ground because the common ground only records information that is shared by the speaker and interlocutor. This indicates that adjustments to the common ground and completed acts of telling will co-vary.

When the speaker presents her interlocutor with a testimonial reason for belief, she is not engaged in the unilateral pursuit of trying to change the interlocutor’s mind in the way that Greco and Lawlor suggest, nor is the interlocutor simply looking to the speaker as a source of evidence. Instead, the speaker and interlocutor are engaged in changing the context between them. Telling is a joint action because it is a negotiation in which the speaker and interlocutor direct their attention to a proposition, and jointly determine how it gets added to the common ground of their exchange.

## **5. Further benefits of this account**

I have argued that telling is a form of joint action because it involves a speaker and interlocutor who jointly determine the common ground of their exchange while in a state of

shared understanding. But what are the results of thinking of telling in this way? In this section, I briefly survey some of the possible benefits of my account.

First, understanding telling as a form of joint action could help us model and explain some of the ways in which others influence what we say. For example, it could help to explain why the practical effects of our speech do not always match up to what we intend. That is, because interlocutors help to not only complete acts of telling but also to determine what, if anything, is added to the common ground, we can explain some of the ways in which our words “get away from us” through reference to the fact that acts of telling are not wholly determined by speakers. In this way, my account could also help to explain ongoing philosophical debates about silencing where speakers from marginalized groups find their efforts to perform certain acts of telling frustrated.<sup>30</sup>

Understanding telling as a form of joint action might also help to explain the many acts of telling that are performed by speakers who are not fully aware of what it is that they want to tell their interlocutors. In some of these cases, the speaker’s message is only apparent after some back and forth with the interlocutor. An example of this is when a child is helped by a caretaker who finishes her thought when she struggles to express herself. In such cases, understanding telling as a form of joint action might help to explain how it is that interlocutors can help speakers determine which proposition they would like to attest to.

Finally, my account of telling as a form of joint action provides a way in which joint actions may occur even in the absence of cooperation of a certain kind. This would expand our understanding of the phenomenon by making a range of new examples available for analysis. After all, as Annette Baier observes, “[i]f we want eventually to get an understanding of all the

---

<sup>30</sup> See e.g., Langton (1993).

intentional multi-person actions and activities [...] we need to keep in view [...] the competitive, mutually interfering and even hostile activities as well as the harmonious and non-competitive ones” (Baier 1997, 27).

## **6. Conclusion**

In this paper, I have defended the view that telling is a form of joint action. I began by arguing that while telling is a form of joint action, it should not be understood in the same way that paradigm cases of joint action are typically understood. To do so, I presented Greco’s argument that the successful acts of telling—cases in which the speaker is in fact believed and knowledge is transmitted—essentially involve joint action. To build his case, Greco draws on the concept of shared intention, which many in the joint action literature believe to be the distinguishing feature of joint action. However, I argued that this is a mistake because shared intentions are only capable of explaining actions that are primarily cooperative and seem incompatible with many situations in which acts of telling regularly occur. This includes arguments and situations that involve a level of antagonism, as well as any situation in which speakers and interlocutors share few overlapping goals or interests.

Instead, I proposed an account of telling as joint action which does not build in a level of cooperation that is incompatible with many acts of telling. To do this, I drew on insights from Moran, Tugendhat, and Stalnaker. From Moran, I drew on the idea that acts of telling involve the giving and receiving of a special, testimonial kind of reason for belief. The giving and receiving of testimonial reasons in turn gives rise to an intersubjective or mutual encounter in which the speaker and interlocutor achieve a shared understanding of the proposition that the speaker attests to. This is evident when we consider Tugendhat’s insight that acts of telling always make

room for a linguistic response from the interlocutor, because it highlights the fact that the speaker and interlocutor are jointly attending to the very same thing—a proposition—that the interlocutor can respond to with affirmation, denial, or abstention. Finally, I drew on Stalnaker’s notion of common ground to show that acts of telling are not merely intersubjective experiences as Moran argues but are simultaneously the joint determination of the common ground of the exchange.

I concluded by briefly surveying some of the benefits of my account. Understanding telling as a form of joint action helps to explain some of the ways that others seem to influence what we say, particularly in situations where speakers and interlocutors have different levels of power and authority. It also shows that joint action occurs in a range of contexts and between participants with different abilities or degrees of cognitive sophistication. While my account represents a departure from the kinds of cases that are focused on in the literature—typically, rational agents performing the kinds of actions over which we think we should have equal jurisdiction—it is one that can positively impact our understanding of the nature of joint action.

## References

- Abbeduto, Leonard, K. Short-Meyerson, G. Benson, and J. Dolish. 2004. "Relationship between theory of mind and language ability in children and adolescents with intellectual disability." *Journal of Intellectual Disability Research* 48 (2): 150-59.
- Alonso, Facundo. 2009. "Shared Intention, Reliance, and Interpersonal Obligation." *Ethics* 119 (3): 444-75.
- Astington, Janet Wilde, and Jennifer M. Jenkins. 1995. "Theory of Mind Development and Social Understanding." *Cognition & Emotion* 9 (2-3): 151-65.
- Austin, J. L. 1962. *How to do Things With Words*. Cambridge: Harvard University Press.
- Baier, Annette C. 1997. "Doing Things With Others: The Mental Commons." In *Commonality and Particularity in Ethics*. Edited by Lilli Alanen, Sara Heinamaa, and Thomas Wallgren, 15-44. London: Macmillan Press Ltd.
- Bratman, Michael E. 1999. *Faces of Intention*. Cambridge: Cambridge University Press.
- Clark, Herbert H. 1996. *Using Language*. Cambridge: Cambridge University Press.
- Fricke, Elizabeth. 2006. "Second-Hand Knowledge." *Philosophy and Phenomenological Research*: 592-618.
- Fricke, Elizabeth. 2012. "Stating and Insinuating." *The Aristotelian Society Supplementary Volume*: 61-94.
- Gelfert, Axel. 2018. "Testimony." *Routledge Encyclopedia of Philosophy*.  
<https://www.rep.routledge.com/articles/thematic/testimony/v-2>.
- Gilbert, Margaret. 1990. "Walking Together: A Paradigmatic Social Phenomenon." *Midwest Studies in Philosophy* 15 (1): 1-14.
- Greco, John. 2020. *The Transmission of Knowledge*. Cambridge: Cambridge University Press.

- Green, Mitchell. 2017. "Conversation and Common Ground." *Philosophical Studies* 174 (6): 1587-1604.
- Grice, H. Paul. 1957. "Meaning." *The Philosophical Review* 66 (3): 377-88.
- Grice, H. Paul. 1975. "Logic and Conversation." In *Syntax and Semantics, Vol. 3, Speech Acts*. Edited by Peter Cole and Jerry L. Morgan, 41-58. New York: Academic Press.
- Langton, Rae. 1993. "Speech Acts and Unspeakable Acts." *Philosophy & Public Affairs* 22 (4): 293-330.
- Lawlor, Krista. 2021. "Telling as Joint Action: Comments on Richard Moran's The Exchange of Words." *Philosophy and Phenomenological* 102 (3): 701-7.
- McMahon, Christopher. 2005. "Shared Agency and Rational Cooperation." *Nous* 39 (2): 284-308.
- Miller, Seumas. 2001. *Social Action: A Teleological Account*. New York: Cambridge University Press.
- Moran, Richard. 2018. *The Exchange of Words: Speech, Testimony, and Intersubjectivity*. Oxford: Oxford University Press.
- Moran, Richard. 2021. "Replies to Lawlor, Goldberg, and Pritchard." *Philosophy and Phenomenological Research* 102 (3): 721-8.
- Roth, Abraham Sesshu. 2017. "Shared Agency." *Stanford Encyclopedia of Philosophy*. Last modified May 1, 2017. <https://plato.stanford.edu/archives/sum2017/entries/shared-agency/>.
- Searle, John. 1990. "Collective Intentions and Actions." In *Intentions in Communication*, edited by Philip R. Cohen, Jerry Morgan and Martha Pollack, 401-15. Cambridge: MIT Press.
- Stalnaker, Robert. 2002. "Common Ground." *Linguistics and Philosophy* 25:701-21.
- Stalnaker, Robert. 2014. *Context*. Oxford: Oxford University Press.



Thomason, Richmond H. 1990. "Accommodation, Meaning, and Implicature: Interdisciplinary Foundations for Pragmatics." In *Intentions in Communication*, edited by Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, 325-63. Cambridge: MIT Press.

Tugendhat, Ernst. 1982. "The Employment Rule of an Assertoric Sentence." In *Traditional and Analytic Philosophy: Lectures on the Philosophy of Language*, by Ernst Tugendhat, 186-200. Cambridge: Cambridge University Press.

## Chapter 3: Telling, Indirection, and Joint Action

### 0. Introduction

As we saw in Chapter 2, some philosophers believe that the paradigmatic case of testimony occurs when a speaker in a face-to-face encounter openly *tells* her interlocutor that *p* and thereby assumes responsibility for the truth of *p*.<sup>31</sup> On this view, the speaker's assumption of responsibility is what makes the transmission of knowledge via testimony possible—it is what licenses or justifies the interlocutor's taking the speaker at her word and coming to know what she knows. But recently, some philosophers (Fricker 2012; Peet 2015; Moran 2018) have noted that this view of telling, with its emphasis on a speaker's intentional and overt statement of some *p* with the express purpose of making her knowledge of *p* available to her interlocutor, is at odds with some of the other ways in which speakers present information to audiences, for example via insinuation, hedging, or irony—in other words, indirect forms of communication.<sup>32</sup>

Generally speaking, indirection is found in cases where what is actually said diverges from what is meant or communicated. A simple remark of: “Oh, that’s what you’re wearing!” may literally express surprise at someone’s choice of attire, but in particular contexts may have the force of outright criticism (“That outfit is inappropriate”) or perhaps a gentler hint that there is some mismatch between one’s clothing and the weather, one’s destination, and so on. Indirect forms of communication such as this, common as they appear to be in ordinary communication, are in tension with telling-based views of testimony. The tension arises because indirect forms of communication seem to allow the speaker to avoid responsibility for the truth of a proposition *p*

---

<sup>31</sup> See e.g., Fricker (2012, 592); Moran (2018); Goldberg (2006); McMyler (2013).

<sup>32</sup> For hedging see Benton and van Elswyk’s (2020) “Hedged Assertions.” For insinuating see Fricker’s (2012) “Stating and Insinuating” and Moran’s (2018) *The Exchange of Words*.

that she indirectly communicates to her interlocutor. And, because defenders of the view above believe that the speaker's assumption of responsibility is part of what facilitates the transmission of testimonial knowledge, it seems that the transmission of testimonial knowledge is incompatible with indirect speech (Ross 1986; Fricker 2006; Moran 2018). To put it another way, because the speaker can deny responsibility for what she implies, the interlocutor lacks the sort of justification he would have if he were to believe what a speaker says in a paradigmatic case of testimony.

Different philosophers have provided different explanations for the speaker's apparent ability to avoid assuming responsibility for what she communicates indirectly. For Elizabeth Fricker (2012), speakers cannot be regarded as having assumed "unambiguous responsibility for the truth of any *particular* message" if there is more than one way to plausibly interpret the speaker's utterance (Fricker 2012, 79-80; my emphasis). Indirection is consistent with multiple interpretations because the implied content is not communicated via a conventional linguistic vehicle which would help to fix a particular message. Instead, indirect speech conveys a "fuzzy equivalence class of messages" (e.g., "I hate that outfit," "I think you'll get cold," or "I don't think you got the memo about the dress code"). Richard Moran (2018) however, believes that indirect speech acts such as insinuation may fall short of paradigmatic tellings not because what the speaker intends with her speech is ambiguous—the interlocutor may be perfectly confident that the speaker intends a particular message (e.g., "I hate that outfit")—but because the speaker fails to make her intention open and avowable between them. The result for both Fricker and Moran is the same: speakers seem able to avoid taking responsibility for what they indirectly communicate.

The general phenomenon of responsibility avoidance in indirect communication is familiar enough. We can easily imagine that the speaker who comments: “Oh, that’s what you’re wearing,” does in fact intend to criticise her interlocutor’s attire but is able to avoid taking responsibility for having criticised because she does not do so openly. The views surveyed above suggest that this form of communication is also incompatible with testimony, a significant source of knowledge in day-to-day life. But if this is true, then a significant source of our knowledge is threatened, because speakers frequently deploy indirection in their speech. In fact, stating things outright in the way that telling-based views of testimony seem to require is relatively uncommon (Tannen 1986, 1989, 23; Lakoff 1979; Brown and Levinson 1987). Fortunately, however, the problem that indirect speech presents for the transmission of testimonial knowledge is often overcome in the course of ordinary conversation. Because while it is true that speakers can and do avoid taking responsibility for content they indirectly communicate, they can sometimes be *held* responsible for that content by their interlocutors. But how does this happen?

Intuitively, indirection presents opportunities for interlocutors to play bigger roles as interpreters of speech than they otherwise would when what is said by a speaker is explicit and clear. Extending this intuition, I argue that at least some forms of indirection involve an even bigger role for the interlocutor; that of disambiguating or making avowable the speech act that is performed. This means that the transmission of knowledge is not undermined when speakers say things indirectly, because the speaker is not solely responsible for determining what speech act is performed. We see this when we take seriously the idea that was defended in the last chapter: that telling is a form of joint action in which the speaker and interlocutor jointly determine the common ground of their exchange. In at least some cases of indirection, the interlocutor helps to constitute the common ground by making the communicated content determinate. This brings the

unstated but intended message out into the open, as it is in the paradigmatic case in which the speaker openly attests to some content and thereby assumes responsibility for its truth. If the interlocutor successfully brings the unstated content out into the open, then problems having to do with the ambiguous or unavowable nature of the message are dissolved, and nothing blocks the speaker's assumption of responsibility and the transmission of knowledge.

I begin with a brief introduction to Fricker and Moran's telling-based views of testimony as well as the general phenomenon of indirect speech. Then, I discuss the tension between telling-based views of testimony and indirection in speech, focusing on Fricker's more developed account. I suggest that we should reject Fricker and Moran's arguments because the prevalence of indirection in ordinary communication would undermine the possibility of testimonial knowledge. I then propose a way of understanding how it is that testimonial knowledge may be transmitted from a speaker to an interlocutor, even in cases where the speaker avoids openly stating what she intends. When we understand telling as a form of joint action in which speakers and interlocutors jointly constitute the speech act that is performed, we see that speakers can be held responsible for that which they do not explicitly state. That is, because it is a joint rather than unilateral action performed by only the speaker, the interlocutor can step in and contribute to speaker's fulfilment of her role in the joint action. The interlocutor does so by making the content that the speaker indirectly communicates open and avowable between them.

## **1. Telling and Testimony**

Telling-based views of testimony seek to explain the particular way in which we gain reasons for belief and knowledge from one another via speech. When one person successfully transmits her knowledge of  $p$  to an audience by telling her audience that  $p$  and being believed,

her audience seems to come to know what she knows. But how does this happen? For some, audiences are licensed to believe what speakers say and come to know what speakers know because in the act of telling, speakers openly take responsibility for the truth of *p*.

For example, Elizabeth Fricker (2006, 2012) claims that the act of telling involves a speaker who openly and intentionally expresses her knowledge of some fact to an audience and thereby makes that same knowledge available to her interlocutor for belief. For Fricker, this public act of taking responsibility for the truth of some *p* and thereby making available to her audience something she knows involves the “socially recognised and maintained performative force” of the speaker giving her audience her word that *p* is the case (2012, 62). In short, telling someone something is governed by various conventions that make it possible for the transmission of knowledge to take place. For Fricker, “crucial amongst the conventional norms governing assertoric tellings is an epistemic norm: one should assert that *P* only if one knows that *P* (the 'K-norm' for assertion)” (62). This norm is crucial for the interlocutor’s belief in *p* on the basis of the speaker’s telling him that *p* because “assertion is governed by the K-norm is part of what enables a recipient of testimony to know that the testifier has that property of trustworthiness - not easily would she tell that *P*, unless she knew” (Fricker, 2012; 63). The speaker who fails to adhere to the K-norm risks social sanction and the interlocutor’s understanding that the speaker is likely to only take responsibility for some *p* if she knows that *p* is part of what licenses the interlocutor’s believing on the basis of what the speaker tells him.<sup>33</sup>

---

<sup>33</sup> This account of what licenses the interlocutor’s belief is an evidential view of telling, according to which an interlocutor is licensed to believe if he has evidence that the speaker is in fact saying what she takes to be true and can generally be relied upon to say true things. Moran (2018) explicitly rejects this characterization of the kind of reasons speakers present to interlocutors when they tell them something.

Along similar lines, Richard Moran (2018) defends an assurance view of telling according to which a speaker tells her interlocutor something by explicitly presenting a reason for belief and thereby assuming responsibility for the truth of what she says. Like Fricker, Moran believes that the speaker's intentional act of telling her interlocutor that *p* plays a key part in her overt or public assumption of responsibility and provides a kind of guarantee to her interlocutor of the truth of what she says. However, unlike Fricker, Moran does not think that the interlocutor's reason for belief is grounded in the empirical evidence that the speaker is likely to only tell people things that are true rather than false because of the threat of social sanction. Instead, Moran claims that part of what constitutes a telling is the set of responsibilities and entitlements that come to exist when a speaker tells someone something and is believed (2018, 5). For Moran, the speaker's public assumption of responsibility has the consequence of making the speaker accountable to her interlocutor for the truth of what she says: the speaker commits herself to the truth of some proposition by presenting it to her interlocutor as true, and in doing so confers various entitlements (e.g., of complaint) on her interlocutor.

Telling-based views of testimony therefore attempt to explain testimonial knowledge in terms of the speech act of telling, and specifically the way in which telling allows a speaker to take responsibility for the truth of *p* when she openly and explicitly attests to *p*. However, as I discuss in the next section, many communicative acts are not explicit in the way that telling-based views of testimony require.

## **2. Indirect speech acts**

In the last section, we saw that some philosophers try to explain how testimonial knowledge is possible or justified by appeal to the fact that in telling—the paradigmatic case of

testimony—the speaker assumes responsibility for the truth of  $p$  with her open and explicit utterance of “ $p$ .” The interlocutor can therefore take the speaker at her word and believe what he is told because of this assumption of responsibility by the speaker, to whom he can complain if  $p$  turns out to be false. However, much of our everyday experience of communicating with others involves broadly indirect forms of speech, which these philosophers believe is inconsistent with telling someone something. Before turning to the problem that indirect speech acts pose for telling-based views of testimony, I first provide a brief overview of indirect speech acts.

In linguistics and the philosophy of language, indirect speech is studied as speech acts in which what is meant diverges from or communicates something over and above the conventional meaning of our utterance.<sup>34</sup> We can begin with John Searle’s (1975) classic example:

(1) Can you pass the salt?

The standard interpretation of this sentence is that the speaker who utters it *requests* that her addressee passes the salt. However, this request is indirect, because what is literally said takes the form of a question about the addressee’s ability to pass the salt. As Searle puts it, indirect speech involves the indirect performance of one illocutionary act (the request) by way of performing another illocutionary act (the question) (1975, 168). For Searle, we can understand utterances like these as cases in which the speaker may both mean what she actually says as well

---

<sup>34</sup> This classic understanding of indirect speech articulated by Searle and others is the standard view but is not universally accepted. Recently, Lepore and Stone (2018) have argued that the indirect content of such speech is explicitly encoded in the grammar of such sentences. Thus, expressions like (1) present “a complex package of related meanings together as a single unit” by *both* raising the question of whether the interlocutor can pass the salt and expressing a request to pass the salt. However, even on this view, Lepore and Stone acknowledge that the presence of multiple meanings only weakly commits the speaker and affords “flexible responses” to the addressee. They also observe that the response the interlocutor happens to choose (e.g., “yes” or “no” in response to (1)) seems to resolve the conversational state in one way or another: if the interlocutor says “yes,” the exchange resolves into a directive to pass the salt with compliance from the interlocutor, while if the interlocutor says “no,” “the speaker is *not* committed to the directive and the addressee has *not* rejected it” (Lepore and Stone, 2018). I therefore leave aside the question of whether indirection is explicitly encoded or pragmatically deduced since both accounts are compatible with the view I go on to defend.



as something else with different propositional content that corresponds to what is indirectly meant. Interlocutors retrieve this indirect content by relying on things like mutually shared background information, as well as their own general powers of rationality and inference (169).

Cases like (1) seem relatively unproblematic as instances of indirection. When speakers request that someone passes the salt by uttering the sentence (1) in ordinary contexts, interlocutors have no problem retrieving the indirect content (in this case, the request).<sup>35</sup> Speakers would also find it hard to plausibly deny, with the utterance of (1) in ordinary contexts, that they did not intend to request that their addressee pass the salt. This is because “Can you pass the salt” is conventionally used by many English speakers to request salt. And, because (1) is conventionally used to request that someone pass the salt, we might say that the request is at least almost as explicit and open as it would be if the speaker said instead: “Please pass the salt.”

So, let us turn to an example in which the communication of the indirect speech act by way of another speech act is both nonconventional and an assertion, rather than a request:

(2) This was a great movie choice, Pauline.

The literal or conventional meaning of this utterance is that Pauline’s choice of movie was great. However, we might fill in this example in a way that makes it clear that the speaker intends to communicate something else entirely. For example, the speaker may have uttered this sentence in a situation where Pauline’s choice has plainly turned out to be a terrible one, or a situation in which some movie trope the speaker despises has repeatedly come up.

---

<sup>35</sup> Similarly, sentences like “x costs an arm and a leg” and “We’ll cross that bridge when we come to it” are conventionally used to communicate things that diverge from the literal meanings of the sentences. If it is agreed that sentences like (1) are examples of indirect speech this shows that at least some cases of indirect speech are unproblematic.

In this sort of case, we might, along with Searle, say that the speaker has performed the illocutionary act of ridiculing Pauline's movie choice indirectly, by way of an illocutionary act of praising Pauline's choice. Or, in Gricean terms, we might say that the speaker has conversationally implicated that Pauline's movie choice was a bad one, by way of saying that it was a great one.<sup>36</sup> In either of these views, Pauline will have to work out the indirect content of the sentence via various context clues and mutually shared background information—for example, the fact that she knows the speaker hates the trope the movie turns out to contain, or the fact that she can infer some sarcasm from the speaker's dry tone, and so on.

This idea that indirect speech acts require that interlocutors figure out a speaker's communicative act by drawing on the context in which the sentence is uttered as well as shared background information (that is, information not included in what the speaker actually states) is key to the problem of indirection identified by Fricker and Moran. I turn to this problem in the next section.

### **3. The problem of indirection for telling-based views of testimony**

I have so far glossed the idea that telling-based views of testimony seem incompatible with indirect speech acts, or speech acts in which a speaker communicates some content *x*, by way of uttering “*y*.” But why think that telling-based views of testimony are incompatible with indirection in general, which seems like a common enough occurrence in our everyday experience of language? In what follows, I present Fricker, Peet, and Moran's arguments for this conclusion.

I begin with Fricker, who summarizes her view thusly:

---

<sup>36</sup> See Grice (1975).

“There are several important dissimilarities between the communicative act of overtly stating that P to an audience and the sneakier one of stating something else, say Q, with the intention of thereby getting them also to infer that P, but without oneself stating P. Only the former [...] constitutes full-strength testifying to the fact that P. (Fricker 2012, 64)

Here, Fricker draws a clear line between paradigmatic tellings in which a speaker overtly attests to some fact by stating that *p* and cases in which speakers only imply that fact indirectly, by stating that *q*. The latter kind of cases involve what Fricker calls “Implicit Secondary Messages,” or “Implicatures” (72). For Fricker, Implicit Secondary Messages (ISM) include “the range of cases where the speaker makes an explicit linguistic statement, an assertion of a primary message, but in addition intends to convey an unstated implication(s) of this” (72).

When a speaker implies but does not state that *p*, her interlocutor must rely on the “knowledge-context” or “mutually known relevant features of context plus other background knowledge” in order to retrieve the speaker’s implied message (73). However, doing so is cognitively burdensome and liable to result in misunderstandings, because retrieving the correct message requires that the speaker and interlocutor coordinate their knowledge context—roughly, they should be on the same page regarding conversational direction, the general situation in which they find themselves, as well as each other’s second-order representations about all of this (87). Since knowledge contexts are unlikely to perfectly overlap, this allows speakers to plausibly maintain that the interlocutor retrieved the wrong message.

Fricker believes that there are two main factors which make it the case that speakers cannot be held accountable for what they conversationally imply but do not state outright. The first has to do with the “ambiguous epistemics” of implicatures (87). To see what Fricker means

by this, it is useful to take a detour through a particular class of communicative acts that Fricker calls “One-Off Grice” (OOG) acts—one-off cases of agential meaning which do not use any conventionally mediated symbols to convey an intended message. Such cases may occur when we communicate through various expressions and gestures that do not fix a particular message, as in the case where we exchange a look with a friend across the room when something surprising happens. In OOG cases, interlocutors must rely fully on the knowledge context to identify the message the speaker intends to convey because speaker does not provide any linguistic clues or conventional signs to aid in this identification. These communicative acts share the problematic ambiguous epistemics of ISM cases.

For Fricker, these OOG cases are significant because recipients will for the most part be unable to retrieve a specific intended message from an agent’s non-conventionally mediated gesture, because this gesture does not fix a specific content. As she puts it, “in the absence of expressions with a conventional semantics, a *specific* message cannot be recovered. Instead, there is a fuzzy equivalence class of messages” (79, Fricker’s emphasis). In OOG cases, the speaker’s gesture does not fix a particular content in the way a paradigmatic telling does, instead there are likely to be several plausible ways in which to interpret the speaker’s act of communication.

As an illustration of Fricker’s claim about OOGs, consider a case in which I walk into a room where there are several people already present and make eye contact with a friend who begins to gesture animatedly at me. We have not previously been in a conversation so there is no linguistic context that I can use to determine what she is trying to communicate to me. So, to determine the content of my friend’s message I must rely on the knowledge context—including what I know about the other people in the room and why they might be there, along with what I

know about my friend and what she is likely to be trying to communicate in this context.

Suppose that I conclude from my friend's gestures that I have walked into the room at a fraught moment—perhaps because one of the occupants is in the middle of recounting a harrowing ordeal—and that I should try to be quiet and not interrupt. Suppose also that my friend appears satisfied with my response to what I take to be her intended message—with my quickly and quietly settling into the room to avoid interrupting the story. That is, let us suppose that it appears as if a successful act of non-verbal communication has taken place. But can we say that a successful transference of testimonial knowledge has taken place?

Fricker thinks not, because of the low likelihood that recipients of OOG acts of communication retrieve a specific message from these acts, rather than a “fuzzy equivalence class” of messages. In our above example, I might retrieve from my friend's gesture the message that “Someone is sharing something important, try not to interrupt by making too much noise,” but it might be that my friend's actual intended message was closer to “Don't be your usual loud self, someone is sharing something important,” or even, “Get in here and listen, this is interesting/surprising/entertaining.” We can imagine iterations of this case in which the existence of this fuzzy equivalence class becomes plain: if I interpret my friend's gestures as the second possible message above and later complain that I didn't need to be told not to be loud, my friend may plausibly claim that her intended message was the first or third message, and she did not mean to suggest I wouldn't have figured out that I needed to be quiet.<sup>37</sup> This suggests that even if the recipient is confident that she has retrieved the correct message, there is very likely an open

---

<sup>37</sup> Or alternatively, that she had no particular message in mind at the time. These cases are treated equivalently by Fricker.

possibility that there are a number of other messages that may be plausibly attributed to the utterer.

For this reason, Fricker concludes that OOG cases fall well short of the paradigmatic case of telling, in which the speaker openly expresses a specific content and thereby takes responsibility for its truth:

[...]in One-Off Grice there is no determination of a specific primary message. This being so, there is no question of the utterer signing up, by her utterance, to responsibility for truth of a specific content.” (Fricker 2012, 76)

Again, because OOG cases do not involve the determination of a specific content by the speaker, but rather a fuzzy equivalence class of messages, the speaker cannot be said to have taken responsibility for the truth of some specific content.

This problem of ambiguous epistemics extends to implicatures because they too involve the issue of the speaker failing to give voice to a specific content. In Fricker’s ISMs, or equivalently in implicatures, or what I am here broadly referring to as indirect speech, speakers explicitly attest to some *p* while intending to convey some further message, *q*. But when a speaker implies but does not say what she means, the interlocutor is then forced to rely heavily on the knowledge-context of the conversation. However, because the speaker and interlocutor’s respective understanding of the knowledge-context is unlikely to perfectly overlap, this opens up a space in which misunderstandings become possible, if not likely. And, because these situations involve a higher possibility of misunderstanding, or failure of coordination, it will always be possible for the speaker to exploit this fact to deny that she intended to convey or commit to a particular *p*. That is, the speaker will always be able to point to an (alleged) error in the interlocutor’s understanding of the knowledge-context to deny that the interlocutor retrieved her

intended message. Thus, cases like this allow the speaker to avoid responsibility for some particular proposition  $p$ , because they can always maintain that they actually meant that  $q$ .

The second factor which allows speakers to avoid responsibility for the things which they conversationally imply is based on a distinction made by Kant. Kant (1963) observes that there is a moral difference between lying and merely bringing it about that someone believes a falsehood via other means. Similarly, Fricker thinks there is an important moral difference between explicitly giving your word that  $p$  is the case and merely allowing an interlocutor to infer that  $p$  is the case. This moral difference between stating and implying as well as the ambiguous epistemics of conversational implicature together explain what Fricker takes to be an observable fact: that people do not treat speakers as having asserted when they have only implied (2012, 86).

In “Testimony, Pragmatics, and Plausible Deniability,” Andrew Peet (2015) suggests that the deniability problem is more far-reaching than Fricker appreciated. While Fricker thinks that the deniability problem is only present in situations where there is some implied secondary messaging that diverges from the primary message in which it is couched, Peet believes that the problem arises in any situation where knowledge context must be deployed to retrieve the meaning of a statement. Thus, Peet claims that the deniability problem extends beyond indirect speech to any proposition that is conveyed using context sensitive expressions.

An example of the basic phenomenon Peet is referring to is found in expressions like: “He’s killed every one of them.” The meaning of the expression is easily retrieved, but retrieving the content requires knowledge of the context in which it was uttered—in this case, knowledge of the fact that the speaker was in the midst of assessing Jack’s poor track record with house plants. Context sensitive expressions abound in our everyday use of language, but they are often easy enough to understand. Very young children, for example, understand that “I’m tired” has a

content that varies because “I” will typically refer to whoever utters the expression. But Peet is suggesting that context sensitivity is at the heart of the deniability problem because it creates an opportunity for the speaker to deny that her audience retrieves the correct content of a context sensitive utterance. And, if context sensitivity is widespread, then speakers are very often in positions where they can deny that audiences retrieve the correct content. More specifically, he suggests that content recovery that relies on “extensive appeal to knowledge context” will give rise to the deniability problem:

“In general, if recovery of an asserted content requires extensive appeal to knowledge context then the speaker will often be able to claim that the audience recovered the wrong proposition, thus disclaiming responsibility for the audience's belief. Therefore, if heavy duty appeal to knowledge context is often required for recovering what is said, and this appeal to knowledge context gives rise to the same possibilities of error to which implicature gives rise, then the deniability problem will apply to a wide range of assertions, not just implicatures.” (Peet 2015, 33)

Note that Peet does not say the audience is likely to retrieve the wrong proposition if extensive appeal to knowledge context is required to retrieve that proposition. Audiences are demonstrably able to retrieve the correct proposition in a vast number of cases—and do so quite easily. Instead, reliance on the knowledge context merely opens up a space in which the speaker has room to construct a more or less plausible story about how the audience retrieved the wrong proposition—even if they did not, and even if the audience interpreted the speaker’s assertion in the most plausible way.



Moran's (2018) concern about indirection or cases in which speakers imply but do not state outright what they intend diverges from those of Fricker and Peet because Moran does not think the problem is one of ambiguity:

“When someone insinuates but does not say that her interlocutor is a liar, that person may be as certain as can be of the speaker's intentions, yet they may remain unavowable between them, as in the case of open hypocrisy. The unavowability of insinuation lies elsewhere.” (Moran 2018, 180)

For Moran, a speech act is avowable with respect to an audience when a speaker can openly announce what act she is performing with her words by being able to tack “I hereby” onto her statement, as in “I hereby tell you it is two o'clock.” Acts that fail this test of avowability include insinuation, because an insinuation is not the kind of thing a speaker openly announces to her interlocutor, and not the kind of act a speaker can tack “I hereby” onto (as in, “I hereby insinuate that you are a liar”), even in cases where it is obvious to them both what the speaker is doing. A speaker may insinuate that her interlocutor is a liar, and it may be clear to them both what the speaker has insinuated, but for Moran, because the speaker holds back from an “outright saying [or telling],” her act remains unavowable in that “she does not present her interlocutor with that statement to reply to” (180). In sum, for Moran, “avowability is [...] central to the notion of *saying something* in the sense of making a claim or statement with one's words” (181).<sup>38</sup>

This idea has some intuitive force. Consider again a case in which a speaker comments: “Oh, that's what you're wearing” to an interlocutor with whom the speaker is attending an event. Even if we build into the case that the speaker intended to insinuate that the interlocutor is

---

<sup>38</sup> “Saying something” as Moran uses is here is equivalent to telling someone something.

inappropriately dressed—perhaps the speaker and interlocutor are a parent and teenager—the parent will in many cases be able to avoid responsibility for what she says by simply failing to make her criticism avowable between them. In fact, it seems like this avoidance of responsibility could occur even though, perhaps due to past critical remarks, it is obvious to the teenager that he was being criticized. For Moran, by failing to make her intention to criticize avowable between them, the speaker also does not provide the interlocutor with statement to which he can reply. And, because Moran believes that acts of telling always make a reply possible, the speaker in cases like these does not tell her interlocutor that his outfit is inappropriate. And, because she does not tell him that his outfit is inappropriate, it seems she cannot be held responsible for having done so.

For Moran then, acts of telling must take the form of statements, that is, they must be clear expressions of the proposition in question because this is how speakers commit themselves to or take responsibility for the truth of particular propositions. In the absence of such an expression, no act of telling takes place.

Fricker, Peet, and Moran therefore agree, for various reasons, that telling someone something is incompatible with indirection, implicature, or speech acts in which, to use Moran's terminology, some conveyed message remains unavowable between the speaker and interlocutor. In the next section, I turn to my argument that the problem identified by Fricker, Moran, and Peet is sometimes overcome by an interlocutor who makes the speaker's intended proposition open and avowable between them.

#### **4. Indirect speech reconsidered**

Before turning to my positive view, I want to make some general critical remarks about the idea that indirection in speech creates the problems discussed above. Recall that, according to the telling-based views of testimony, the paradigmatic form of telling involves a speaker who openly and explicitly tells her interlocutor that *p*. The speaker's open declaration of *p* is important for telling-based views of testimony because an open declaration is the means by which speakers commit themselves to the truth of some *p*; it is how they give their interlocutors their word that *p* is true. The speaker's open commitment to the truth of *p*, or equivalently, her assumption of responsibility for the truth of *p*, in turn makes it possible for her to transmit her knowledge of *p* to the interlocutor. This is because the speaker's act of taking responsibility for the truth of *p* is what justifies the interlocutor's belief in what she says, and what facilitates the interlocutor's coming to know what the speaker knows.

Because telling-based views of testimony rely on this structure to explain how the transmission of testimonial knowledge is possible, these views of testimony seem incompatible with indirect speech. In indirect speech, the speaker does not openly express that some *p* is true and therefore does not take responsibility for the truth of *p*. This seems to block the transmission of knowledge that would otherwise take place if the speaker had expressed herself openly or unambiguously.

But it is not clear that we can so easily distinguish between direct and indirect speech acts. Consider the following example from Grice (1975, 44): "He is in the grip of a vice." As Grice observes, in order to determine whether the speaker who utters this sentence means that the subject of the sentence is unable to rid himself of a bad habit or has been caught in a clamping tool, we would need to know more about the context in which it is uttered. That is, we would need to know the identity of the subject of the sentence and which meaning of "in the grip of a

vice” the speaker intended. This is significant because it seems like this ambiguity opens up a space for a speaker to deny, for example, that she was speculating about the personal struggles of the subject of the sentence and construct a more or less plausible story about an entirely different person who had an accident with a clamping tool. However, “He is in the grip of a vice” is not an indirect expression of what the speaker meant, and so it seems that the problematic features of indirect speech are not limited to indirect speech.

A similar issue is raised by phrases like “I went to the bank today.” While there are two ways of interpreting bank, and thus some opportunity for speakers to try to wriggle out of taking responsibility for a particular content, the expression itself seems perfectly clear and direct because the conventional meaning of the expression is likely closely tied to the intended meaning, even though there are multiple conventional meanings of the word “bank.” This is part of what I take to be Peet’s point in his discussion of Fricker’s argument. Because Fricker suggests that the availability of multiple interpretations of a given phrase can allow speakers to avoid fully committing to a specific one, the problem potentially generalizes beyond the indirect speech acts that Fricker discusses, because there are sometimes multiple ways to interpret expressions that are more or less direct.

But even if we were to grant that there is some principled way to distinguish between direct and indirect speech, it is not clear that this distinction tracks with the contents for which speakers can be held responsible. The main reason for this is that indirection abounds in ordinary communication, but rarely presents problems for interlocutors who must retrieve the intended content and come to know what the speaker knows. In fact, as Tannen points out, we are constantly tracking indirect, non-linguistic information about how speakers express themselves even when what they say is clear and direct (1986, 66). Many of us have been faced with

someone who tensely grits out that “It’s fine” or “I’m not upset” even as their body language communicates the opposite. In these cases, it is not unusual for an interlocutor to press the issue with something like “You don’t seem fine,” or “Why did you say it like that then?” because they have noticed the mismatch between the speaker’s body language and explicit message. How should we understand examples like this?

If we went along with the division between paradigmatic acts of telling and indirect forms of communication proposed by Fricker and Moran, we would have to conclude that speaker in such cases is responsible for the truth of her claim that she is not angry unless she explicitly states otherwise. However, this conclusion sits uneasily with what goes on in many such cases. Suppose that Greta and Mia are friends who have the sort of exchange described above. In such a case, the exchange might go like this:

Mia: -tensely- “I’m fine”

Greta: “Are you sure? You don’t seem fine...”

Mia: -pauses, visibly deflates- “I just would have liked to have been invited”

Here, it would be strange to insist that the Mia is responsible for the truth of her claim that she is fine, even though she does not explicitly denounce that initial claim or say that she is not fine. Instead, Greta’s ability to extract information from not only what Mia said but how she said it has made Mia’s being upset at being excluded mutually acknowledged and open between them.<sup>39</sup>

---

<sup>39</sup> Here, Moran might protest that what Greta learns about Mia from her tone or body language is not strictly speaking testimonial knowledge, but instead an evidential reason that Greta derives on her own. I don’t disagree with Moran’s distinction; however, I believe that we oversimplify how communication actually works when we divorce evidential reasons from testimonial ones. In the example, Greta does not approach Mia in the way a detective would; instead, she uses the extra-linguistic evidence to encourage Mia to confide in her and remove the distance created by Mia’s initial attempt to disavow being upset.

Tannen believes that we constantly monitor these non-linguistic “metamessages” because they contain information about our relationships to one another (1986, 66). In the present example, Greta picks up on the fact that her friend is angry and trying to distance herself and responds by inviting Mia to confide in her. Thus, it seems that interlocutors regularly tailor their responses to reflect both the direct, linguistic information as well as the indirect metamessages that are here conveyed extra-linguistically. Our monitoring of these metamessages (even in cases of ostensibly “direct” communication) suggests, as Tannen argues, that “indirectness, in the sense of metamessages, is basic to communication” (1986, 66).

We also easily retrieve metamessages that are communicated via linguistic vehicles. For example, Jeanne, a teenager, may try to determine whether her parents are amenable to her going to a party without straightforwardly asking to go to that party. As Tannen points out, if someone in Jeanne’s position did straightforwardly ask to go to the party, she would face having her request be denied as well as the conflict and confrontation that would follow. So instead, Jeanne may try to raise the issue of going to the party without making a request, perhaps by saying things like “Oliver—I think you know his parents—is having a party on Friday night” or alternatively “All my friends are going to this party on Friday night” and so on. Despite her indirect approach, we can easily imagine that Jeanne’s parents would have no trouble retrieving the indirect content of their daughter’s words. This is especially clear in cases where the parents choose to respond to that content, rather than what Jeanne directly states. For example, they may say “Yes, you can go” or “You’ve gone out twice this month already!” Of course, if her parents respond with the latter, Jeanne may try to protest that she was not asking, but such a protest need not successfully absolve her of responsibility. In many cases, Jeanne will accept that her parents,

who know her very well, have worked out that she wanted to go to the party and have denied her request.

The prevalence of cases like these, as well as indirection in general, suggests that we should reject Fricker and Moran's claim that indirection is inconsistent with the transmission of knowledge. That is, because indirection seems to be, as Tannen suggests, a basic part of speech and communication, the conclusion that it is incompatible with the transmission of knowledge would undermine much of what we come to know on the basis of what we are told by others. Given that we do seem to successfully engage in the transmission of knowledge, we must be able to do so despite the fact that indirection is prevalent in our communicative endeavours. I provide an account of how this occurs in the next section.

## **5. Indirection and joint action**

Consider the following case described by Tannen (1986):

Cynthia told Greg she was hurt because he fixed himself a snack without offering her any. So he offered her the snack he had just fixed. She turned it down. He asked why. Because he hadn't prepared it for her. Greg was exasperated. Was she hungry or not?

To Cynthia, whether or not she was hungry was beside the point; the point was whether or not Greg thought about her when he fixed himself a snack, which showed whether or not he cared about her as much as she cared about him. She would never feed herself without asking him "Would you like some?" In fact, she might not even have a snack if he didn't want one. (Tannen 1986, 66)

Why doesn't Cynthia simply tell Greg that she wants him to show that he cares by thinking of her when he fixes a snack for himself? As Tannen points out, being direct in this case would undermine what Cynthia wants, which is for Greg to consider her in the way that she considers him without being prompted. If Cynthia did come out and say "I want you to show me that you care about me" or something similar, and Greg responded "I do care," it is unlikely that Cynthia would feel any better than she did after Greg failed to ask her if she wanted a snack. The fact that she had to force him to say that he cares will feel hollow to her, thus being direct in this situation would have been counterproductive.

This example further highlights the fact that there is much that is missing from Fricker and Moran's characterization of indirect speech. Specifically, Fricker and Moran both focus on speakers who use indirection to get away with something by maintaining plausible deniability. In such cases, speakers state that  $p$  while implying  $q$ , but deny that the content ( $q$ ) retrieved by their interlocutor was content they intended, and they are able to do so because  $q$  is not stated outright, but presented in such a way that it is deniable, or disavowable, later on. This may be done by a speaker who is from the start of the exchange aiming to encode her message in such a way that it is deniable or by a speaker who only aims to do so after the fact, perhaps because the content retrieved has given some offense or not been well received.

But not every indirect speech act is performed by a speaker who wishes to avoid taking responsibility for what they intend to convey. As Fricker acknowledges, speakers may also use indirection because it is quicker. This typically occurs in cases where the speaker knows that her interlocutor can be relied upon to work out what she means even with relatively few "utterance clues," or clues the speaker includes in her utterance to ensure that her message is recoverable by her intended audience (2012, 67). Thus, a speaker may simply say to a close friend who has



stopped by: “There’s a bottle of wine in the kitchen,” while leaving it to the friend to infer that he is welcome to help himself. In some such cases, the fact that the speaker does not need to fully spell out her intended message also helps to reaffirm the relationship that exists between the speaker and the interlocutor. In fact, the interlocutor may understandably be taken aback if a close friend formally offered to pour him a glass of wine.

In still other cases, indirect speech may be an invitation to an interlocutor to work out a puzzle. This might occur when the speaker aims to be witty or intends to make a joke by saying one thing and meaning another. For example, the speaker in (2) might be lightly poking fun at Pauline for her movie choice by saying “This was a great movie choice, Pauline.” In a situation where Pauline is aware that the speaker despises the tropes that turned out to be abundantly featured in the movie, she will have to work out why the speaker is saying something that is obviously false. In this case, assuming that the speaker and Pauline are well acquainted, she will work out that the speaker means to poke fun at her for picking a movie that the speaker does not enjoy.

When interlocutors successfully work out the puzzle left for them by an interlocutor, it helps to build or reaffirm the relationship between them. Thus, in response to a friend’s query about what we should eat for dinner, I may say something like “Well, I do like pizza” to which my friend may say something like “Classic Pizza?” In this exchange, more is communicated than is directly stated, and if I responded to a similar query from someone whom I did not know very well, they might be understandably mystified by my response. However, because my friend and I have what Searle described as “mutually shared background information,” my friend was able to easily understand my apparently unhelpful response. We can even push this example further by supposing that when my friend asks, “What should we eat?” and I merely raise my eyebrows and

incline my head towards the drawer where we keep the menu for Classic Pizza. Given enough shared history or mutually shared background information, my friend will easily understand what I mean, and this exchange will have reaffirmed our friendship.

These cases are important because although they are not quite the kinds of cases that Fricker and Moran have in mind, they demonstrate that indirect speech is deployed for a variety of reasons, and speakers understand that interlocutors can and regularly do retrieve content that is communicated indirectly. When we notice this fact, it is easier to see how interlocutors might wield this ability to accurately retrieve indirect content when faced with a speaker who, in bad faith, tries to disavow or deny responsibility for that indirect content. And, when interlocutors successfully bring this indirect content out into the open, the issues of ambiguity or unavailability that make the speaker's avoidance of responsibility possible are dissolved, thus clearing a path for the interlocutor to hold the speaker responsible for that indirect content.<sup>40</sup>

Consider a case in which a worker is trying to get out of working late for the fourth day in a row. Because her boss can be temperamental, she is not sure if requesting to go home and avoid working late will be met with displeasure or some form of punishment. She therefore eases herself into the difficult conversation by trying to avoid openly stating what she wants. She says things like, "Oh, this would be the fourth day in a row" or "My family is beginning to forget what I look like." Here, as the previous examples suggest, her boss is very likely to figure out her intended meaning. Thus, he will pick up on the fact that the worker would like to refuse his request to stay late.

---

<sup>40</sup> I am here glossing over some difficult details about how this happens. For example, it isn't clear when and how interlocutors successfully make the indirect content open, and when and how speakers can block this. These are deep questions that I do not have the space to explore here, so instead I gesture to what I take to be one successful instance of an interlocutor bringing the indirect content out into the open and holding the interlocutor responsible for it.

At this stage, we cannot yet say that the worker can be held responsible for trying to refuse to work late. While the boss might be perfectly confident that the worker wants to refuse his request to stay late, it is not yet open or avowable between them, as something that they mutually understand or entertain. As we saw in Chapter 2, this distinction matters because it seems to mark a difference between two people whose actions happen to coincide but act individually and two people who are engaged in a mutual encounter. Fricker appears to identify the same issue with the speaker's privately held intentions:

Since the claim that she intended to communicate E turns on claims about her private intentions and [knowledge context] including her second order representations of others' beliefs, she can always get away with denying that she intended any such thing; even if her denial is made in bad faith. Lies about my own intentions and other mental states may be suspected, but cannot be refuted. In contrast, when someone makes an explicit statement of a fact P, what she signs up to in doing so – taking responsibility for the truth of P – is a public fact about the situation, determined by semantics and objective features of context. So it cannot be incorrigibly denied by the speaker. I can be nailed as having stated that P; never as having insinuated that P. (Fricker 2012, 88-9)

Fricker, however, does not consider that the interlocutor might be able to make the intention in question explicit. In the current case, the boss may do so by denying the worker's unstated request to go home by saying (perhaps): "No, I actually need you to work late today."

What has gone on here? Again, it seems that despite the fact that the worker does not directly state her request, her boss seems to easily retrieve the unstated proposition. That is, we can understand him as responding not to what she more or less explicitly tells him—that her

family thinks she is working too much—but to a request to go home. This is why he seems to *deny* a request to go home. According to Fricker, speakers in situations like this maintain plausible deniability about what proposition they are in fact committed to, and so the boss cannot hold the worker responsible for having tried to refuse. However, I think we can easily imagine that this is not how this particular scenario plays out. That is, despite the fact that she does not openly request to go home early, the worker may find that she faces the same social sanctions she would have faced if she were to have explicitly refused the boss's request. In other words, the worker might be held responsible for something that she does not openly express.

How is it possible for the worker to be held responsible for something that she does not openly express? The answer, I believe, has to do with the fact that telling is a form of joint action. As discussed in Chapter 2, this is not because the worker and the boss share an intention to pursue a shared goal—quite the opposite. In the scenario above, the worker and the boss are at odds: she would like to go home to her family while he would like to increase productivity. Yet despite their differences, the boss and the worker are able to coordinate on the indirect content of her speech. He figures out that she would like to go home early, even though she does not explicitly request to go home early. He also seems to deny her request, even though her request is never explicitly stated.

In openly denying her implied request, the boss effectively brings into the open and makes precise the worker's indirect refusal. This simultaneously makes it the case that the worker and the boss share an understanding of her unstated request to go home early. And, in a situation in which the worker cannot or does not protest and is treated poorly after the fact because of this request, we can also conclude that this proposition was entered into the conversational record, or the common ground of their exchange. In cases like these, the

interlocutor is responsible for not only recognizing the speaker's intention and helping to determine whether and how the proposition she expresses is added to the common ground, but also for retrieving a specific proposition from what the speaker says and does not say.

Note that this view does not entail that the interlocutor can usurp the speaker's authority or unilaterally determine what speech act is in fact performed. It does entail that the speaker does not have *sole* authority over the determination of what speech act is performed, particularly in indirect cases. Again, this is not because sole authority over what act is performed shifts from the speaker to the interlocutor, but because part of what will inform the interlocutor's retrieval of the relevant, unstated proposition is the context between them and the various linguistic and extra-linguistic clues made available by the speaker. And, part of what will make it a legitimate or accurate retrieval is that it conforms in some way to what the speaker actually intended.

## **6. Conclusion**

In this chapter, I have argued that the problem of indirect speech for telling-based views of testimony is sometimes overcome by interlocutors who disambiguate and otherwise make explicit the speaker's intended but unstated message. I began by introducing Fricker and Moran's telling-based views of testimony. According to these views, speakers transmit their knowledge of some proposition  $p$  to their interlocutors by openly and explicitly asserting that  $p$  is the case. The speaker's explicit assertion of a particular proposition is important because this is how speakers assume responsibility for the truth of what they say, thus enabling their interlocutors to believe what they say, or come to know what they know.

Next, I discussed the phenomenon of indirect speech which involves content that is implied, but not explicitly stated. According to Fricker, Peet, and Moran, indirect speech poses a

problem for the transmission of knowledge, because it allows speakers to avoid responsibility for what they say. For Fricker and Peet, this is because indirect speech makes the speaker's commitment ambiguous: she cannot be held responsible for a particular proposition because she maintains plausible deniability about the exact proposition to which she is committed. For Moran, indirect speech acts like insinuation pose a problem for the transmission of knowledge not because they are ambiguous but because they are not avowable. For Moran, because insinuated content is not avowable, the speaker does not provide the interlocutor with a statement to which he can respond, which in turn does not make the content open between them. When a speaker communicates a message that is not avowable to her interlocutor, she does not make an overt claim that the (unavowed) message is true, thus she does not take responsibility for that content. In both views, because the speaker cannot be held responsible for what she implicates or insinuates, she cannot transmit knowledge of the insinuated content to the interlocutor.

I then turned to my positive view, in which I first tried to show that the distinction between direct and indirect speech acts is not as clear as Fricker and Moran suggest. I also suggested that the problems of ambiguity identified by Fricker are not limited to indirect forms of speech. However, I argued that this is not a problem for speakers and interlocutors in much of their everyday communicative exchanges. Most competent users of language will easily pick up on indirect information that is communicated by speakers both linguistically and extra-linguistically. This challenges Fricker's view that interlocutors cannot retrieve specific messages when these messages are communicated indirectly. I also suggested that, contra Moran, when interlocutors do retrieve these messages, they are able to make them avowable by expressing them openly even though the speaker fails to do so. This means that the speakers' ability to avoid responsibility for what they imply but do not state is held in check. In some cases, the

interlocutor's input can push the speaker to commit to the implied content. In other cases, like that of the worker and her boss, the interlocutor can instead make the speaker's intended message open and avowable between them. This is possible because the speaker is not solely responsible for determining what speech act she performs: instead, she shares that responsibility with the interlocutor to whom she directs her utterance.

In sum, while speakers can and do sometimes avoid taking responsibility for what they indirectly communicate to their interlocutors, they can also be held responsible by interlocutors who work to retrieve and make explicit the intentions that speakers do not state outright. What determines whether speakers successfully avoid responsibility for what they imply but do not state? While it is not clear to me that there is a principled way of determining whether speakers will avoid responsibility in a given case, the examples discussed above suggest that the answer has much to do with the relative power or authority wielded by the speaker and interlocutor. A boss can hold his worker to something she merely implies in much the same way that a parent can hold a child responsible. It is less likely that the reverse is true.

## References

- Benton, Matthew A., and Peter van Elswyk. 2020. "Hedged Assertion." In *The Oxford Handbook of Assertion*, edited by Sanford Goldberg, 245-63. Oxford: Oxford University Press.
- Brown, Penelope, and Stephen C. Levinson. 1987. *Politeness: Some Universals in Language Use*. Cambridge: Cambridge University Press.
- Clark, Herbert H. 1996. *Using Language*. Cambridge: Cambridge University Press.
- Fricke, Elizabeth. 2006. "Second-Hand Knowledge." *Philosophy and Phenomenological Research* 73 (3): 592-618.
- Fricke, Elizabeth. 2012. "Stating and Insinuating." *Proceedings of the Aristotelian Society, Supplementary Volumes* 86 (1): 61-94.
- Goldberg, Sanford C. 2006. "Brown on Self-Knowledge and Discriminability." *Pacific Philosophical Quarterly* 87 (3): 301-14.
- Grice, H.P. 1957. "Meaning." *The Philosophical Review* 66 (3): 377-388.
- Grice, H.P. 1975. "Logic and Conversation." In *Syntax and Semantics, Vol. 3, Speech Acts*, edited by Peter Cole and Jerry L. Morgan, 41-58. New York: Academic Press.
- Kant, Immanuel. 1963. *Lectures on Ethics*. Translated by Louis Infield. New York: Harper and Row.
- Lakoff, Robin Tolmach. 1979. "Stylistic Strategies Within a Grammar of Style." *Language, Sex, and Gender: Does La Difference Make a Difference* 327 (1): 53-78.
- Lepore, Ernie, and Matthew Stone. 2018. "Explicit Indirection." In *New Work on Speech Acts*, edited by Daniel Fogal, Daniel W. Harris, and Matt Moss. Oxford: Oxford University Press.
- McMyler, Benjamin. 2013. "The Epistemic Significance of Address." *Synthese* 190 (6): 1059-78.



- Moran, Richard. 2018. *The Exchange of Words: Speech, Testimony, and Intersubjectivity*. New York: Oxford University Press.
- Peet, Andrew. 2015. "Testimony, Pragmatics, and Plausible Deniability." *Episteme* 12 (1): 29-51.
- Ross, Angus. 1986. "Why Do We Believe What We Are Told?" *Ratio* 28 (1): 69-88.
- Searle, John. 1975. "Indirect Speech Acts." In *Syntax and Semantics*, edited by Peter Cole and Jerry L. Morgan, 59-82. New York: Academic Press.
- Stalnaker, Robert. 2014. *Context*. Oxford: Oxford University Press.
- Tannen, Deborah. 1986. *That's Not What I Meant!* New York: HarperCollins Publishers.
- Tannen, Deborah. 1989. *Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse*. Cambridge: Cambridge University Press.
- Tugendhat, Ernst. 1982. "The Employment Rule of an Assertoric Sentence." In *Traditional and Analytic Philosophy: Lectures on the Philosophy of Language*, by Ernst Tugendhat, 186-200. Cambridge: Cambridge University Press.

## Chapter 4: Developmental Joint Action

### 0. Introduction

In a 1975 study on early conversational interactions between mothers and infants,<sup>41</sup> Mary Catherine Bateson describes how one three-month-old infant, “Mackie,” appears to respond to his mother as she chats to him after a bath. In response to each of his mother’s questions: “You gonna be a good boy today? [...] You’re not? [...] What? [...] You are?!” Mackie makes soft, cooing vocalizations that are coordinated with his mother’s speech (Bateson 1975). In addition, and like the other one-to-three-month-olds in the study, Mackie rarely interrupts when it is his mother’s turn to speak, appearing to time his vocalizations so that they harmonize but do not overlap with those of his mother.

Bateson’s findings have since been corroborated by numerous studies, including several other pioneering studies from the early 1970s in which researchers used video to facilitate detailed microanalysis of infant behaviour.<sup>42</sup> These early studies constituted a revolution in modern psychology. Up until that point, pre-verbal infants were still regularly thought of as asocial and egocentric, armed only with a set of instinctive behaviours that could attract caretakers to meet their biological needs.<sup>43</sup> Aided by advancements in film which led to the possibility of video-based studies and frame-by-frame microanalysis, researchers were able to challenge these old assumptions with new evidence of the pre-verbal infant’s rich interpersonal world. In addition to the “proto-conversations” that Baby Mackie engages in with his mother, researchers have found that at just several weeks old, infants engage in a range of reciprocal interactions with their adult caretakers. These include mutual gazing, reciprocal imitation,

---

<sup>41</sup> Throughout, I use “infant” to refer to children under a year old.

<sup>42</sup> See Bateson (1975, 1979), Stern (1971, 1974, 1985, 2000), and Trevarthen (1979).

<sup>43</sup> See Trevarthen (2011b, 120; 2011a, 86), Rochat (2001), and Reddy (2008).

emotion-sharing, and making anticipatory body adjustments to enable the action of being picked up or fed (Trevarthen 1979; Reddy, Markova, and Wallot 2013).

In this chapter, I will argue that we have good reason to conceive of some of these early interactions, which can occur before the age of six months, as a form of joint action. This form of joint action occurs primarily between the infant and the infant's primary adult caretakers and departs in many ways from leading accounts of joint action in philosophy, which tend to focus on joint action between adults. It also departs from some recent and important accounts of joint action in young children.<sup>44</sup> These accounts, which were developed to address the leading accounts' apparent exclusion of young children, generally aim to relax the cognitive requirements of many leading accounts. But while these alternative accounts of joint action do not require the same level of cognitive sophistication, they typically focus on children who are at least twelve to eighteen months of age and have acquired the capacity to engage in child-friendly analogues of paradigm forms of joint action (i.e., building a tower together rather than making a Bolognese sauce together).

This form of joint action is, to use a term from developmental psychologists, *triadic*, because it involves a child and adult who act with respect to some third object, event, or fact which is external to them and to which they jointly direct their attention. Importantly, this triadic structure is not in place when the child and the adult merely attend to, for example, building a tower at the same time; instead, they must additionally be aware that the other so attends (Tomasello 1995). The joint state of attending to the building of the tower then underpins and helps to coordinate their joint activity of building the tower together. For psychologists like Michael Tomasello, the ability to engage in these sorts of interactions shows that the child

---

<sup>44</sup> See e.g., Tollefsen (2005); Butterfill (2012); Pacherie (2013); Story (2021).

already understands the other with whom she jointly attends and acts as a subject of experiences and pursuer of goals that may coincide with or diverge from her own (1995, 107). Thus, the child can recognize that she shares a common interest with another and play her part in simple collaborative projects related to that interest.

The early interactions that I will focus on in this paper are, by contrast, *dyadic*, because they involve face-to-face interactions between an infant and caretaker who attend only to each other and do not coordinate with respect to some external object of interest. These interactions occur during the stage of development that psychologists call primary intersubjectivity, which spans the period from birth (or by some more conservative estimates, two months) to around nine months of age, at which point infants begin to enter the stage of triadic, secondary intersubjectivity, which is firmly in place by the time they are around twelve months old.<sup>45</sup> While triadic interactions slot more easily into paradigm forms of joint action in the philosophical literature, which typically involve two adult agents who jointly pursue some common goal or project like painting a house together or singing a duet, I do not believe they exhaust the infant's capacity for acting jointly with others.

Instead, I will argue that early dyadic interactions represent a form of joint action that has been neglected in the philosophical literature. This form of joint action does not involve identifying that another individual has a common interest in some external object which may be jointly pursued, but it does involve intersubjective experiences which serve to create, maintain, or negotiate social bonds. These bonds are arguably an important part of why we engage in certain kinds of joint action (like spending time with others) even as adults, yet it is often

---

<sup>45</sup> Colwyn Trevarthen (1979) was the first to call this stage "primary intersubjectivity."

downplayed in accounts that focus on the instrumental or productive form of joint action, in which individuals coordinate to bring about a shared goal.<sup>46</sup> Focusing on the stage of primary intersubjectivity therefore allows us to explore this underappreciated aspect of joint action and is potentially instructive for our understanding of forms of joint action that we continue to engage in as adults. As I will argue, all that is needed for this form of joint action is coordinated action and the ability to share experiences, which in turn forms a kind of common ground between participants.<sup>47</sup>

I begin in the next section with an overview of some empirical evidence of the infant's ability to engage in a variety of mutual and reciprocal interactions with her primary adult caretakers and highlight why some developmental psychologists believe they are genuinely shared experiences. I then discuss several alternative accounts of joint action recently put forward by philosophers and consider whether these accounts can explain the reciprocal interactions engaged in by infants at the very earliest stages of life. I argue that they cannot do so, in part because they focus on the triadic, external-goal-oriented forms of joint action that have become standard in the literature. I then defend the idea that these early reciprocal

---

<sup>46</sup> Notably, even in cases where spending time with others is centred around some sort of activity like watching a movie or playing a board game, there are plausibly multiple joint actions in play, one of which is simply spending time in the company of others who also want to spend time with you.

<sup>47</sup> There is disagreement among developmental psychologists about whether primary intersubjectivity constitutes a genuinely shared experience. See Tomasello et al. (2005); Moll and Tomasello (2010). Tomasello believes that these early interactions are not shared in the relevant sense because they are not triadic while Reddy (2008, 2011; Brinck, Reddy, and Zahavi 2017) believes that they are shared because they are triadic. I do not have the space to give a full account of this disagreement, but I will broadly side with those like Reddy (2008) and Trevarthen (1979, 2011) who believe that these early experiences are indeed intersubjective and shared. In contrast to Reddy, however, I will not claim that these interactions are triadic. Instead, I suggest that no external object is needed for a genuinely intersubjective experience. See Moll et al. (2021) for a discussion of the dispute between Reddy and Tomasello.

interactions between infants and their caretakers are, despite their departure from paradigm cases, a form of joint action.

### **1. Empirical data: evidence and discussion**

For several decades now, researchers have been gathering evidence that shows that newborns are attuned to other people, especially their primary caretakers, in several ways. For example, it is well-established that infants are sensitive and attracted to human faces, preferring to look at even schematic faces over other patterns drawn on head-sized discs (Fantz 1963). This sensitivity appears innate, as Colwyn Trevarthen observes: “Within minutes of birth, an alert newborn may orient to the face and eyes of a happy mother” and “recognition of her face can be acquired within hours of birth” (2011a, 83).<sup>48</sup> We also find evidence of the infant’s attunement to faces in perturbation experiments. In one such experiment, researchers asked the mother of a seven-day-old infant to put on a ski mask as she went about her normal routine of caregiving (Sander 2007, 187-88). At first, the infant appeared not to notice that anything unusual was afoot, and the mother went about her caregiving routine normally. However, when it was time for a feeding and the mother positioned the infant so that they were face-to-face, the infant reacted with surprise and the infant’s previously calm state was transformed into a disturbed and unsettled one for the duration of the feeding.

Infants also display a particular interest in eyes and strongly prefer to look at faces with open, rather than closed eyes (Batki et al. 2000). In one study, researchers found that when presented with photographs of faces that either looked directly at them or faces that looked away, infants preferred by a large margin those images with faces that appeared to make direct eye

---

<sup>48</sup> See also Bushnell, Sai, and Mullin (1989).

contact over those with the faces with averted eyes (Farroni et al., 2002). This effect also occurred when infants were presented with schematic faces but disappeared if the features on the face were scrambled or if the face was upside down (Farroni et al., 2006).

The infant's attunement to human faces and eyes makes several reciprocal interactions possible. When an alert infant seeks and makes eye contact with an attentive caretaker, it may progress to mutual gazing, which researchers use as a measure of social engagement. This is because eye contact, or at least eye contact of a certain kind, is used by humans to signal a desire to establish a communicative link, and it generally carries positive value.<sup>49</sup> Infants appear sensitive to this and seem to take pleasure in mutual gazing. They are more likely to smile at adult partners who make direct, face-to-face eye contact with them (Csibra 2010, 145). They are also more likely to follow the gaze of an adult who first makes direct eye contact before looking away (Farroni, Johnson, and Csibra 2004). However, when interacting with adult caretakers or experimenters who refuse or prematurely break off eye contact, infants appear to disengage, smiling less often or even avoiding future eye contact with that individual (Trevarthen 2011a, 83). All of this suggests that infants take pleasure in mutual gazing and do not simply seek out faces and eyes because they are exercising an innate ability to detect them. Instead, they do so because they are interested in them and the mutual experiences they facilitate (Csibra 2010, 145).

This general orientation towards human faces and eyes also seems to facilitate infant engagement with facial expressions and movements that communicate affect or emotions and other states of mind (Rochat, Passos-Ferreira, and Salem 2009, 176). When alert and comfortable (i.e., not hungry or tired or otherwise distressed or distracted) some newborns appear to respond

---

<sup>49</sup> Of course, not all eye contact is positive or perceived as positive, and sustained eye contact on many occasions communicates an aim to intimidate.

to the emotional states of others when they are communicated through facial expressions (e.g., smiles or frowns) (Trevvarthen 2011a; Reddy 2008). In fact, infant-caretaker dyads have been observed to mutually regulate affect intensity. Daniel Stern (1998) observes that “in smiling interactions the dyad can increase by increments the level of intensity of the affect [...] one partner increases a smile's intensity eliciting an even bigger smile from the other partner, which ups the level yet again, and so on, producing a positive feedback spiral” (102).

This process increases in sophistication at around six to eight weeks old, when infants begin to use these expressions to elicit a response in kind from their partners, which suggests that they expect and seek out certain reactions in response to certain expressions. This idea has been supported by “still-face” experiments, in which adult caretakers or experimenters suddenly freeze and do not react (i.e., fail to provide the response the infant tries to elicit) while holding the gaze of the infant. In response, infants often appear distressed and frown or look away, or they may attempt to re-engage the frozen caretaker (Rochat, Passos-Ferreira, and Salem 2009, 182; Cohn and Tronick, 1983; Tronick and Cohn, 1989).

From birth, infants are equipped with the capacity to engage in several coordinated, non-reflexive movements. Newborns engage in “pre-reaching,” which involves the coordination of the arm and hand in grasping movements, and they regulate intraoral sucking pressure when bottle feeding (Trevvarthen 2011a). This early coordination forms the basis of the infant’s eventual ability to coordinate her bodily movements with the actions of her caretakers. For example, one study found that infants as young as two months old would adjust their body postures to enable the action of being picked up by their mothers (Reddy, Markova, and Wallot, 2013). As their mothers approached, arms outstretched, the infants anticipated being the action of being picked up by lifting or spreading their arms away from their torsos and then either lifted,



extended, or tucked their legs under their bodies. These movements all served to stiffen the body and facilitate the action of being picked up. In addition, the researchers noted that as the infants made these adjustments, their gaze was largely focused on their mother's faces, which they believed suggested that infants perceived their anticipatory adjustments as their own actions, rather than as something that merely happened to them.<sup>50</sup>

The infant's ability to engage in these simple but coordinated non-reflexive movements is also at play when she imitates her adult caretakers. Though it is disputed, some studies seem to show that infants can imitate an adult's mouth openings, tongue protrusions, or hand movements just minutes after birth (Trevarthen 2011a, 81; Meltzoff and Moore 1977; 1983).<sup>51</sup> Some researchers have also found evidence that neonates (ranging in age from three to fifty-four hours) invite imitation, specifically tongue protrusion gestures (Nagy and Molnar 2004; Nagy 2011). That is, even newborns do not merely copy the movements made by another body, but also appear capable of "actively seeking and engaging with another human being in sympathetic dialog, taking initiative to contribute to its development" (Trevarthen 2011a, 83). This suggests that infants are not just imitating adults to learn particular movements or forms of expression but are instead doing so to reap the emotional benefit of engaging in these two-way, reciprocal interactions (Trevarthen 2011a, 82).

Finally, in the early 1970s, researchers began using film to gather evidence that infants as young as two or three months old were engaging in what Bateson (1975, 1979) termed "proto-conversations." Aided by their ability to perform frame-by-frame microanalysis of the

---

<sup>50</sup> Further evidence that infants perceive these adjustments as their own actions is found in studies that show that infants respond to delays in being picked up by decreasing these anticipatory adjustments. See Fantasia et al. (2016) for discussion.

<sup>51</sup> See Vincini et al. (2017) for discussion of the dispute.

recordings, researchers observed mothers “greeting their infants as partners in dialog, and the infants responded with appropriate timing and changing emotion to their mothers’ delicately patterned expressions of intention and feeling they saw, heard, and felt” (Trevarthen 2011a, 86). In one recording of a premature infant at thirty-two weeks gestational age, the infant’s father begins to imitate her soft cooing vocalizations as she “kangaroos” inside his shirt. The infant then coordinates her vocalizations with those of her father with well-timed responses before smiling in apparent appreciation of their exchange (Trevarthen 2011a, 84).

Evidence that exchanges of this kind are not automatic, as some earlier researchers assumed, is found in studies like Bateson’s, where close attention was paid to the reciprocal nature of these exchanges and the rhythmic synchronization between mother and child (Bateson 1975, 101). Statistical analysis of the exchange between Mackie and his mother that I discussed at the beginning of this chapter revealed that they were vocalizing in turn rather than at random (Bateson 1975, 104). Analysis of the timing of the respective vocalizations of Mackie and his mother also adds credence to the idea that the exchange was a mutual one. Again, when Bateson investigated the idea that the mother was simply inserting vocalizations into a string of random vocalizations by Mackie, she found instead that Mackie would only vocalize at regular intervals. In addition, it was possible to distinguish between a response time and an elicitation time in the vocalizations of both Mackie and his mother. That is, both Mackie and his mother would pause for longer period of time after their own vocalizations in order to *elicit* a response from the other. However, when it was their turn to *respond* to a vocalization made by the other, they did so more quickly (Bateson 1975, 105).

In sum, infants are capable of a range of what appear to be mutual, reciprocal interactions from birth. This is significant not only because it shows that even pre-verbal infants can engage

others along several dimensions, but also because these interactions are either precursors to or the very earliest instances of some of the interpersonal experiences in which we continue to engage throughout our adult lives (Trevarthen and Aitken 2001, 6). The infant's early attunement to others' faces and eyes, as well as her interest in mutual gazing and her ability to track and match the emotions of others, for example, are present in adulthood and can serve the same social bonding purposes at both stages of life. Eye contact is important in face-to-face communication as it can be used to help interlocutors orient towards one another and ensure that they are both engaged in the exchange. Similarly, the infant's distress at eye-contact avoidance seems related to the adult's general distrust of excessive eye-contact avoidance, and the tendency to interpret it as evidence that the other is hiding something or is generally disinterested in communication or engagement. The ability to track and match others' emotions is also a crucial part of our ability to empathize or sympathize with others.

Imitation also serves several interpersonal functions throughout human life. It is readily observable in our everyday interactions that we adopt the postures of others: we may mimic the other's folded arms, or unconsciously jiggle the same leg, or sit up or slouch to match the other. These basic, often unconscious forms of imitation (often called "social mirroring") are present from the very beginning of life and serve the purpose of reinforcing social bonds in various ways. This is perhaps because, as Vasudevi Reddy observes, "imitating [...] seems to be a simple and very powerful way of confirming the other" and can be a "powerful [...] statement of interest [and] connection" (Reddy 2008, 65).<sup>52</sup> For Reddy, this interpersonal confirmation is perhaps why imitation has also been found to be an effective way of connecting with non-verbal

---

<sup>52</sup> Imitation can also be used in other ways of course—a classic playground function of imitation is mockery. However, we seem to engage in the unconscious mirroring form of imitation much more often and this form of imitation generally has a positive valence.

autistic children.

The infant's early communicative efforts are also a crucial part of later communicative abilities. Bateson (1975) remarks that early interactions like those recorded between Baby Mackie and his mother are "proto-conversations" not because they adhere to the more familiar linguistic features of communication but because they adhere to the extra-linguistic regularities that govern conversation. Mackie and his mother, along with the numerous other infant-mother dyads whose exchanges have been analysed in this way, are engaging in the rhythmic aspects of conversation. They follow the regularities governing the "initiation and termination of conversation, alternation and interruption, pacing and interspersing of verbal and nonverbal elements" (Bateson 1975, 110). In effect, Bateson's point is that conversation, even among linguistically competent adults, involves more than just exchanges of information. Instead, we follow rhythms and patterns in our conversational exchanges and often endeavour to match our conversational partners when we do so. This is clear if we consider those times when conversations are uncomfortable, filled with awkward silences or interruptions that impede conversational "flow." The exchanges described above show that infants engage in similar conversational patterns and rhythms from a very young age and are similarly distressed by disruptions.

The similarities between infant and adult interpersonal experiences discussed above is important because it shows that there is a continuity in what we do at the earliest stages and later stages of life. This suggests that infant and adult interpersonal interactions are of a kind, and there is little reason to deny that infants partake in similar interpersonal interactions, despite their relative physical and mental immaturity. My eventual argument will be that these interactions constitute a legitimate form of joint action. That is, they are not forms of action in which infants

engage on their way to traditional forms of joint action; instead, they are fully fledged forms of joint action that can help to explain certain forms or aspects of joint action that we continue to engage in as adults. Before turning to that argument, I consider some recent accounts of joint action involving children. I argue that these accounts, for reasons discussed below, cannot provide a full explanation of these early reciprocal interactions.

## **2. Philosophical accounts of joint action in young children**

While there has been substantial philosophical discussion of joint action in recent decades, relatively little attention has been paid to the phenomenon of joint action in young children. Present accounts of joint action seem to assume that participants are invariably adults, or at least people with the capacity to engage in the usual examples of moving a piece of furniture together, making a Bolognese sauce, taking a trip together, or dancing a tango. However, it is plausible that young children and others who lack the capacity to make a Bolognese sauce or dance a tango still engage in some forms of joint action. Indeed, understanding how children engage in joint action would tell us something about the earliest forms or ontogenetic origins of joint action, which would enrich our understanding of the phenomenon.

In this section, I discuss and evaluate several recent philosophical accounts of joint action in children. While these accounts make important progress in advancing our understanding of early forms of joint action, I do not believe that they manage to capture its ontogenetic origins. This is in part because these accounts confine themselves to the triadic structure which is found in standard cases of joint actions in adults. However, as I will argue in Section 3, the reciprocal

interactions discussed in Section 1, while lacking this triadic structure, nevertheless constitute a form of joint action.

### **2.1 The problem faced by leading accounts of joint action**

Recently, several philosophers have put forward accounts of joint action in young children. These accounts were motivated by the fact that leading accounts of joint action seem to face a deep problem when it comes to this group. The problem, which I briefly glossed in Chapter 2, begins with the observation that leading accounts of joint action often describe the phenomenon as taking place between participants who form complex interlocking intentions to support their joint pursuit of a shared goal and have common knowledge of these intentions. For example, in Michael Bratman's influential account, we share the intention needed to jointly perform some action 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 1(a), 1(b), and meshing subplans of 1(a) and 1(b); you intend that we J in accordance with and because of 1(a), 1(b), and meshing subplans of 1(a) and 1(b).
3. 1 and 2 are common knowledge between us. (Bratman 1999, 121)

On this view, we must not only each intend that we J but also intend that we J in part because we recognize that our potential partner also intends to J with us. In addition, we must intend to ensure that our individual subplans (broadly, the means by which we intend to secure J) mesh or are consistent. For Bratman, this ensures that our activity is means-end rational (1999, 121).

The problem here is that accounts such as this, which rely on interlocking intentions to help organize the actions of two agents around a shared goal, rule out the possibility of joint

action in children under the age of five, because they lack the capacity to form such complex intentions (Tollefsen 2005, 76; Story 2019, 5010-11).<sup>53</sup> However, children under the age of five do seem to engage in joint action. In fact, some psychologists, most notably Michael Tomasello and his colleagues, believe that the earliest forms of joint action play an important role in cognitive development, “including language learning, [...] imitative learning, theory of mind and collaborative action” (Moll et al. 2021).<sup>54</sup> So, it seems likely that leading accounts cannot be the *full* story of joint action, because they cannot explain joint action in which children participate.<sup>55</sup>

To solve this problem, some philosophers have put forward alternative accounts of joint action that are less cognitively demanding to explain how children seem to readily engage in joint actions like rolling a ball back and forth or building a tower together. In the next section, I discuss and assess three accounts which represent three different approaches to solving the problem. While these accounts make important progress in solving the problem that is created by the standard accounts, they do not go far enough, and retain elements of the standard view that needlessly rule out the possibility of joint action at earlier stages of development.

## 2.2 Tollefsen’s alternative account

---

<sup>53</sup> There is evidence that children only develop a robust theory of mind around the age of four or five years old (see Wellman 2002). With it comes the ability to understand, represent, and form intentions about the mental states of others, or in other words, form the complex intentions that are a feature of leading accounts of joint action.

<sup>54</sup> See e.g., Tomasello (1995), Tomasello et al. (2005), and Moll and Tomasello (2010) for discussion of the idea that joint action plays an important role in early cognitive development.

<sup>55</sup> This latter point is Butterfill’s (2012) criticism of paradigm accounts of joint action. Rather than focusing on the age at which young children acquire various mental capacities, which is to some degree variable, Butterfill focuses on the idea that joint action is a key part of developing a theory of mind, and thus theory of mind cannot be what explains joint action in children.

Deborah Tollefsen (2005) argues that the simple games of pretend (e.g., a tea party) that children begin to engage in at around twelve to eighteen months old are a form of joint action. As Tollefsen points out, children at this age who engage in joint pretense at this age are demonstrably able to coordinate their behaviours in ways that support and facilitate the actions of others. However, standard accounts of joint action like Bratman's rule out this candidate example of joint action because they require a robust "theory of mind." Tollefsen takes such a robust theory of mind to involve the following:

- 1) an understanding of other persons in terms of their thoughts, intentions, and beliefs;<sup>56</sup>
- 2) an understanding that other persons' thoughts, beliefs, and intentions may differ from one's own; and
- 3) an understanding that others have thoughts and beliefs that may not match with the current state of affairs (false beliefs). (Tollefsen 2005, 81)

Because Bratman's account requires participants to be mutually responsive to and have common knowledge of the interlocking intentions of their co-participants, it seems that his account requires that participants meet at least the first two conditions. However, since researchers believe that children do not develop a robust theory of mind before the age of four or five years old, Tollefsen argues that Bratman's account, at least in its original form, cannot explain the reality of joint action in children.

To address this gap, Tollefsen proposes a revision of Bratman's account which includes child-friendly analogues of the mutual responsiveness and common knowledge requirements. In

---

<sup>56</sup> Here, Tollefsen is referring to an understanding that other people have thoughts, intentions, and beliefs which facilitates thoughts like "You want to throw the ball."



place of the mutual responsiveness condition, which would require children to be able to understand and respond to the intentions of their co-participants, Tollefsen includes the requirement of being responsive to intentions-in-action. Intentions-in-action broadly refers to the intentions at work in actions that are carried out without a prior intention (Searle 1983, 52-53). The concept is supposed to capture the large number of everyday cases of intentional action that do not involve prior deliberation. In these cases, prior intentions do not cause a later action; instead, the intention-in-action causes the action as it occurs, by representing its conditions of satisfaction (Searle 1983, 53).

To show that young children are responsive to intentions-in-action, Tollefsen cites research that shows that young children can distinguish between outwardly identical behaviours if they can see what occurs prior to these target behaviours (2005, 89). For example, one study showed that twenty-four-month-olds would interpret one set of actions as “trying to open the box” if they saw an adult pulling in various ways at a box before opening it, but not do so if they failed to see the initial pulling sequence. Another study showed that eighteen-month-olds are capable of understanding and carrying out to completion the failed actions of adults that they observed. This suggests that children can understand the general aim or intention of the action, even though they are not yet at the stage where they can reliably identify and track the mental states of other people.

In place of the second problematic requirement of Bratman’s account—the common knowledge requirement—Tollefsen includes a joint attention condition. As discussed earlier, joint attention is thought to develop when children are around nine months old and it coincides with a general ability to engage in a variety of triadic behaviours, or behaviours that involve the child, the other, and some third object of interest. Again, joint attention does not describe what

happens when the child and other both attend to the same object. Instead, they must also be aware that the other is attending to the same object (85). In joint attention, the child can track the objects of attention of the other and can also invite the other to attend to the object of her own interest (84). Importantly, for Tollefsen, this process does not require the child to have knowledge of the mental states of those with whom she jointly attends. Instead, joint attention “simply requires that participants perceive that the other is perceiving,” which children between the ages of twelve to eighteen months seem to readily do (92).

For Tollefsen, these modifications help to secure some of the same functions as Bratman’s original conditions but do so in a way that does not rule out the possibility of joint action in children. Tollefsen’s proposed revision to Bratman’s conditions for shared intention is the following:

- I. (a) I intend that we J and (b) you intend that we J.
- II. I intend that we J in accordance with and because of 1(a) and 1(b), and meshing of subplans of 1a and 1b; you intend the same.
- III. 1 and 2 are jointly perceived. (Tollefsen 2005, 93)

For Tollefsen, this account can explain joint action in children because it does not require the ability to form representations about the contents of other minds. Instead, it simply requires an ability to form intentions and understand the intentions-in-action of the other. Further, because intentions-in-action, unlike intentions, are “perceptually overt” the child is also able to perceive that the relevant activity is shared (93).

But even if Tollefsen provides a plausible revision of Bratman’s account that accommodates joint action in children who are around twelve months old, her account rules out the possibility of joint action in infants. This is because Tollefsen requires a capacity to engage

in specifically triadic joint attention, where in addition to the child and another person, a third object (e.g., a tower) which is jointly attended to is involved. But neonates and infants under six months old do not engage in triadic joint attention (Moll et al. 2021). In their earliest mutual encounters, there is only the child and primary caretaker who attend to each other and react in reciprocal ways to the actions and emotions of each other.

In fact, Tollefsen explicitly denies the possibility of these earlier forms of joint action, claiming that “prior to the first year, young infants are like *windowless monads*” (my emphasis) but that “as they pass the 1-year mark, children begin to engage in more social behavior—in short, more cooperative behavior” (2005, 80). Here, Tollefsen not only implausibly claims that children under a year old are asocial, but also equates social and cooperative behavior. This seems incorrect. It would, after all, be strange if children were more asocial in infancy, the stage when they are generally most reliant on their caretakers for their basic needs and survival. Instead, as the empirical research in Section 1 suggests, children are social in a variety of ways from birth and continue to develop abilities to connect with others *before* they become more independent and less reliant on others. But this does not mean that they cooperate, particularly if we understand cooperation as a social phenomenon wherein two or more people pursue an external end together. Instead, the social life of the infant takes a different form.

### **2.3 Butterfill’s alternative account**

According to Butterfill (2012), Bratman’s account of joint action cannot explain joint action in young children because it presupposes the understanding of other minds that joint action in young children is supposed to explain (28). That is, while psychologists have claimed that engaging in joint action plays a role in explaining how children come to have the abilities to

use concepts like “belief, desire, intention, knowledge and perception, and in the development of higher forms of cognition more generally,” Bratman’s account requires these to be able to share an intention (24). Butterfill endeavours to provide an account of joint action that is compatible with the idea that joint action can help children develop an understanding of other minds.

Like Tollefsen, Butterfill believes that a major issue in Bratman’s account is the requirement that participants form intentions about intentions, but unlike Tollefsen, he chooses to do away with intentions entirely. Guided by a list of features of joint action which he takes to be neither sufficient nor necessary indicators of joint action, Butterfill proposes a simple account of joint action that only requires a capacity to share goals. Butterfill’s conditions for sharing a goal are as follows:

a. *one goal, two or more agents* —

There is a single goal, G, to which each agent’s actions are, or will be, individually directed;

b. *identification* —

each agent can identify each of the other agents in a way that doesn’t depend on knowledge of the goal or actions directed to it;

c. *expectations about goal-directed actions* —

on balance each agent expects each of the other agents she can identify to perform an action directed to the goal; and

d. *expectations about a common effect* —

on balance each agent expects this goal to occur as a common effect of all of their actions directed to the goal, her own and the others’. (Butterfill 2012, 40)

On this view, young children are engaged in joint action if they share a single goal, can identify the other with whom they share a goal (which helps to rule out cases in which we accidentally act with those with whom we happen to share a goal), expect that each participant will contribute, and expect that each participant's contributions will together help to secure the shared goal. For Butterfill, meeting these conditions is sufficient for two or more people to share a goal, and actions driven by a shared goal is enough to conclude that these people are engaged in joint action (44). Further, since children who lack an understanding of other minds may nevertheless share goals and form expectations about goal-related activities and the effects of their actions and the actions of their co-participants, this account explains how children engage in joint action (40).

Unlike Tollefsen, Butterfill does not explicitly rule out the possibility of joint action in infants who cannot yet engage in joint attention and other triadic interactions. Butterfill's account only requires that participants in the joint action can have goals, can identify that others have the same goals, and can form expectations about goal-directed actions and the effects of these goal-directed actions. Can infants under the age of nine to twelve months old meet these conditions?

It does appear that even neonates have goals of, for example, making communicative contact with a primary caretaker (who has the same goal), so it is plausible that neonates can meet the first condition. Similarly, as discussed earlier, neonates can come to recognize primary caretakers within hours of birth and so are plausibly also able to meet Butterfill's second condition. Infants also seem to form expectations about how their adult partners in early reciprocal interactions should behave: perturbation experiments show that infants become distressed or try to re-engage adult partners who suddenly become unresponsive to their communicative efforts. It is less likely, however, that infants could meet Butterfill's fourth

condition because it would mean that an infant who is engaged in proto-conversation, for example, expects the communicative exchange to occur as a common effect of her actions (her vocalizations) as well as the actions of the adult caretaker (speech, “motherese,” etc.). But this already introduces more cognitive sophistication than the infant has at her disposal. Even if the infant already expects the adult to act in certain ways, she does not yet have the capacity to represent the communicative exchange itself as a common effect of her and her caretaker’s actions. This would minimally require a capacity for reasoning about causes and effects, which emerges along with other triadic behaviours at around nine to twelve months old (Moll et al. 2021; Trauble and Pauen 2011).

In fact, the general problem with Butterfill’s conditions is that they are plainly aimed at explaining actions with a triadic structure. Thus, his account requires that participants represent an external goal or fact and represent certain actions as directed towards this goal. This would already entail an ability to engage in means-end reasoning about that goal. However, infants under the age of nine months lack the ability to coordinate with another person around an external goal or object of interest, so Butterfill’s account cannot explain the jointness of the early reciprocal interactions described in Section 1 (Moll et al. 2021, 4).

#### **2.4 Story’s alternative account**

Daniel Story (2021) similarly identifies and tries to address the problem faced by leading accounts of joint action such as Bratman’s. Unlike Tollefsen and Butterfill, however, Story does not attempt to offer a new, less cognitively demanding set of sufficient conditions for joint action involving children. Instead, Story argues that children may act directly on the (shared) intentions

issued to them by an adult (5010). To do so, Story draws on Abraham Roth's (2004) concept of practical intimacy:

**Practical Intimacy:** It is possible for one individual to take up and to act on an intention formed by another without re-issuing the latter's intention. (Roth 2004, 383)

As Roth persuasively argues, practical intimacy in general is not an unusual occurrence: agents regularly adopt the intentions of other agents and do so without deliberation. For example, if you and I are taking a walk together, I might decide to adopt your intentions about which paths we take without devoting any thought to whether we should take these paths, trusting you to select a route that is amenable to me.

How does this happen? For Roth, it is possible for you to issue an intention that we *J* (where *J*-ing in this case is taking a particular path or set of paths) that settles the question of *whether to J* for the both of us. Note that this does not necessarily involve one agent acting against her will or insubordinately in any way, because as Roth points out, part of what legitimizes your settling the question for both of us is that I allow you to do so; that is, I take your settling of the question of whether to *J* to settle the question for us both (2014, 641). Further, Roth believes that when you settle the question of which path to take for us both, I am also entitled to your reasons for taking one particular path over the other, even though I do not consider or come up with these reasons myself. Consequently, these reasons can explain and guide my actions and can also be used to assess whether and how well the intended action was executed (2014, 644).

Story believes that this phenomenon can help explain how young children (twelve-eighteen months) who lack a robust theory of mind might nevertheless engage in joint action.

Roughly, Story's idea is that the cognitive sophistication required by accounts like Bratman's can be offloaded onto the adult in the adult-child pair, though the child can still be an active participant in the joint action. This is at least plausible because, as Roth points out, children regularly act on the intentions of others (2004, 385). It is also not unusual for adults to structure activities in which children engage, or otherwise settle questions about what children do. Thus, in the case of joint actions like making cookies together, Story argues that although the adult may be almost totally responsible for structuring the activity and formulating the relevant intentions, the child can nevertheless be a full participant when she acts directly on the relevant intentions and is thereby entitled to the adult's reasons for so intending (2021, 5021).

Story's account is plausible, and I am sympathetic to his extension of Roth's idea that agents can act directly on the intentions of others. If Story is right, then we would have an account of joint action that explains how actions can be joint despite a deep asymmetry in the contributions of the participants. It also explains the way in which young children seem to engage in joint action—that is, spontaneously and without negotiation or forethought. However, it is unlikely that Story's account could explain the reciprocal interactions described in Section 1. This is because Roth limits the entitlements we have to another person's reasons in practical intimacy to those reasons that we can understand:

“Of course, I am not always so entitled. I might not understand the further reasons for the intention, or the significance of the larger course of events of which my doings are a part. [...] If there is no possibility of communication with those who formed the intention that would somehow allow me to secure the comprehension or understanding necessary for the possibility of endorsement, then that would seem to preclude my



entitlement to those reasons. If there were no way for me to simulate comprehendingly the reasoning of those who have decided so as to capture the further reasons for what I'm doing, then that too would challenge an entitlement to those reasons.” (Roth 2014, 646)

The claim here is that acting on the intentions of another involves entitlement to their practical deliberation—their further reasons for so intending—but that one can only be so entitled if they could understand these reasons, or “the significance of the larger course of events” of which their actions are a part. If this is true, then infants and young children engaged in joint action would need to understand the further reasons behind the intentions on which they act. But it is not clear that even twelve-to-eighteen-month-old children could understand the further reasons behind the some of the intentions on which they seem to readily act when they participate in joint actions like baking cookies together. But even if they did, it is less likely that infants could understand the reasoning behind the intentions formed by their adult caretakers, because this would require a level of cognitive sophistication they lack.

However, it is also unlikely that adult caretakers engage in some of the reciprocal behaviours described in Section 1 after deliberating and forming an intention to, for example, bond with the infant. That is, while it is true that the adult caretaker bonds with the infant when she engages in mutual gazing, affect exchanges, and proto conversation, the caretaker is usually described as engaging in these behaviours naturally, without deliberating or formulating reasons to do so (Trevarthen 2011b). But this is not true of every reciprocal interaction between the infant and the adult caretaker. Being picked up, for example, is a reciprocal interaction to which the infant contributes by stiffening her torso and extending her arms and legs. But it is unlikely that the infant could understand the further reasons behind the adult caretaker's intentional action

of picking her up (which might include getting the infant changed, bathed, fed, etc.). Again, this would require the infant to understand the other as a subject who pursues a goal—in this case picking up the infant—that may coincide with or diverge from the infant’s goals. However, as we have seen, this would involve the ability to engage in triadic experiences, which only begins to develop when the infant is around nine months old (Tomasello 1999, 60).

### **3. Must genuinely intersubjective experiences be triadic?**

In the last section, I presented three accounts of joint action involving young children. These accounts of joint action were developed in response to a problem that seems to be generated by accounts of joint action like Bratman’s which require participants to form complex, interlocking intentions to meet the sufficient conditions for joint action. The problem was that although children lack the robust theory of mind that would allow them to form the complex intentions about intentions required by Bratman’s account, they seem to engage in joint action. The alternative accounts therefore attempt, in various ways, to relax or circumvent the cognitive sophistication required by Bratman-style accounts of shared intention to explain joint action in children.

However, as I have tried to show, these accounts do not go far enough, because they are modeled on actions with the same triadic structure that we see in accounts like Bratman’s. And because this triadic structure already presupposes a certain level of cognitive sophistication, including the ability to represent others as having interests or goals that may coincide with or differ from one’s own, as well as the ability to engage in means-end reasoning, any account that assumes this structure rules out the possibility of joint action in infants who do not yet have these abilities.

One might object that this argument begins with a mistaken premise. That is, one might disagree that the reciprocal interactions in Section 1 constitute joint action and so dispute my claim that accounts like those of Tollefsen, Butterfill and Story fail to capture an earlier form of joint action. At best, the objection might continue, infants under the age of nine months engage in a form of proto-joint action, or a form of emotion-sharing that isn't the same as the more robust sharing of experiences involved in joint action.

This is the view of Tomasello, whose work inspired the philosophical accounts described above. Tomasello believes that the origins of genuinely shared or intersubjective experiences coincides with the onset of joint attentional behaviours and an emerging interest in objects outside the infant-caretaker dyad. This is in turn caused by the development of a “socially recursive mind-reading mechanism” that allows the young child to grasp that the other intends for her to share and jointly pursue a goal (e.g., “I understand that you intend that we J”) (Tomasello et al. 2005). According to Tomasello then, shared experiences are triadic, and the infant-caretaker dyad precedes the infant’s ability to engage in the kind of recursive mind-reading processes that are a hallmark of shared experiences.<sup>57</sup> Henrike Moll, a frequent collaborator of Tomasello, notes that these processes “form the cognitive basis of shared intentionality, including joint attention, cooperative communication, and similar cooperatively structured interactions” (Moll et al. 2021, 2; Tomasello et al. 2005). Thus, in Tomasello’s view, the reciprocal interactions that take place during primary intersubjectivity do not constitute joint action.

Why does Tomasello insist genuinely shared experiences require a triadic structure? In Tomasello’s work on the emergence of shared intentionality, he observes that children begin to

---

<sup>57</sup> See Tomasello et al. (2005); Tomasello and Moll (2010); Tomasello (2018, 2019)

understand that others are goal-oriented at around nine to twelve months old (2005, 682). This capacity allows children to identify that other people sometimes have the same goals as them, and that these goals can be shared and jointly pursued. Importantly, this shared goal between the child and adult “serves to coordinate their activities around the same object triadically and thereby to enable each participant to know something about what the other is perceiving and to predict what she will do next” (Tomasello et al. 2005, 682). In this way, the triadic pursuit of a shared goal tracks the child’s ability to mind-read—that is, to discern what others are perceiving and intending. For these reasons, it is the triadic structure that brings with it “truly intersubjective sharing,” or two people sharing an experience of the same thing at the same time while “knowing together that they are doing this” (Tomasello and Carpenter 2007, 121).

The problem with the dyadic relation between an infant and caretaker is therefore that it lacks a third object which can help the infant or the caretaker know or understand what the other is perceiving or intending. The implication is that without an external object of interest that you and I happen to share, we cannot come together and share an experience in the requisite way. While the adult caretaker will have mind-reading abilities and will often know when the infant is showing an interest in some object, the infant lacks this ability, so there is no common topic or goal between them, and thus no possibility of shared intentionality.

However, I believe we have good reasons to resist this characterization of what counts as a genuinely shared or intersubjective experience. First, Tomasello seems to assume that there is a gulf between our minds and the minds of other people and that the work of the infant is to steadily develop the cognitive capacities required to traverse this gulf. But why accept that there is such a gulf in the first place? When I am walking down the street and a stranger who is travelling in the opposite direction makes eye contact and smiles sunnily at me and I smile back,

we have a kind of exchange that consists solely of the pleasant interaction we have just had. This seems to involve at least a brief meeting of minds, but no external object was needed to facilitate that meeting. It is also plausible, given the continuity between infant and adult experiences, that the infant engages in exchanges very much like this one when her caretaker smiles and she smiles back.<sup>58</sup> However, it does seem plausible that whatever we think of the problem, the empirical data that researchers have gathered about the infant's ability to bond with others shows that the infant is well-equipped to understand and share in the emotional content of other minds, even before she is able to jointly attend to external objects of interest.

Another reason to resist Tomasello's characterization of genuinely shared experiences has to do with the so-called mind-reading processes that help the infant understand other minds. As Richard Moran (2018) convincingly argues, socially recursive mind-reading can sometimes fail to secure what we might think of as a genuinely social experience. Suppose, for example, that I would like you to know about something significant that has happened in my life, perhaps a falling out with a family member. However, I do not want to talk about it with you, and so instead I leave you with hints or clues that help you deduce what I would like you to know. If you do come to know what I want you to know, you will have done so via the socially recursive mind-reading process that Tomasello describes: you will come to know that I intend for you to know that I had a falling out with my family member. However, you could do this more or less independently of me. You might find and decode my clues and hints in the same way a detective might decode the clues left at a crime scene. In a case like this, you come to know about my intentional state without having engaged in a social or mutual encounter with me. The onset of

---

<sup>58</sup> Of course, part of what Tomasello is expressing is simply the general problem of other minds that has occupied philosophers for thousands of years, and I make no attempt to try to solve it here.

mind-reading abilities might therefore be understood as the onset of the capacity to understand people at a distance.

If we accept that there are good reasons to reject the view that shared experiences are necessarily triadic, then we are well-placed to see why early dyadic interactions are genuinely shared experiences and a form of joint action. I now turn to a defense of this view.

#### **4. The ontogeny of joint action**

My claim in this chapter is that the reciprocal interactions in which infants and their caretakers engage constitute a form of joint action, and that this form of joint action is structurally different from the forms of joint action discussed in both standard and alternative, children-accommodating accounts of joint action. In the last section, I suggested that the triadic structure, in which two or more participants attend to and pursue some external goal, is unnecessary. In this section, I offer an account of how these dyadic interactions might be thought to amount to joint actions.

As a reminder, the dyadic, reciprocal interactions at issue here include mutual gazing, reciprocal imitation, emotion-sharing, carrying and feeding, and proto-conversations. As we saw in Section 1, these interactions are reciprocal because infants seem to willingly contribute to them when they are in a calm and alert state, and further, seem to take pleasure in the fact that the adult caretaker also contributes. Still, these interactions plainly do not involve shared intentionality, nor do they conform to the accounts of joint action in children discussed in Section 2. So, why think that these early interactions are a form of joint action?

I suggest that this form of joint action requires only the ability to:

- i. Share intersubjective experiences

- ii. Coordinate one's actions with the actions of those with whom you share the experience

Taken together, these conditions help to explain how it is that infant-caretaker dyads engage in joint action despite lacking an external object towards which they co-orient. The first condition requires an ability to share experiences. By this, I mean that participants should be able to experience something—whether it is a feeling or an action—while another person experiences the same feeling or action in a potentially slightly different way, and they both attend to each other. Shared experiences are related to what I referred to as shared understanding in Chapters 2 and 3, however, where shared understanding requires that two people attend to the very same proposition, sharing experiences only requires that participants attend to each other while they are engaged in the shared experience. As discussed earlier, infants in dyadic interactions attend to their adult counterparts, as these counterparts attend to them.

The second condition adds that the participants in the joint action should coordinate their actions with those of their co-experiencer. We find ample evidence of the infant's ability to coordinate in this way with her adult caretaker when she adjusts her body in a way that facilitates the action of being picked up, or when she rhythmically responds to her caretaker's vocalizations. The infant's actions, even at this early age, can be responsive to the actions of those with whom she interacts. This is also supported by the fact that infants seem to form expectations about the actions of the other when they are engaged in reciprocal interactions. As we saw earlier, the infant either becomes upset or appears to try to re-engage adult counterparts who suddenly interrupt or cut short these reciprocal interactions.

While very minimal, I believe these proposed conditions secure the two main aspects of any form of joint action—both inner, or mental, coordination as well as external coordination.

When an alert and calm infant engages in mutual gazing with her adult caretaker and smiles to respond to or elicit a smile, she has met the requirements for engaging in this form of joint action. This analysis can also help to explain some forms of action that we continue to engage in even as adults. Some straightforward cases of this include the example of strangers smiling at each other in the street, as well as the conversational rhythms and regularities of non-linguistic proto-conversation that we continue to engage in even after we learn to engage in linguistic communication.

There is in fact a useful comparison to be made between these two broad aspects of conversation and what might be called the two broad aspects of joint action. As Bateson points out, while conversation is often construed as an exchange of information, this construal leaves out the importance of the flow of conversation; the rhythmic, turn-taking aspect of it. Similarly, in much of the current philosophical literature, including the alternative accounts discussed above, the focus is often on the instrumental or productive element of joint action. Agents coordinate their actions with one another in order to pursue a shared goal, and analyses of this kind of activity in turn focus on the most efficient or rational way to achieve these goals. This structure remains even in the alternative accounts, where the goals are rather more frivolous, and often occur in the context of play. This is in one way unsurprising. If we begin with the notion that joint action is for coordinating our activity so that we may achieve a shared goal, we end up with accounts of joint action that are like the ones discussed above.

But while this is no doubt an important feature of at least some forms of joint action, it may not be the only one. As Schmid (2014) has forcefully argued, this approach ignores the intersubjective experience that is also a crucial feature of joint action. In this way, many extant accounts make a mistake when they assume that the phenomenon of joint action is really about



the cooperative and productive effort to bring about an external goal. As we see clearly when we engage in any form of joint action that lacks a clear external goal, like spending time with loved ones, or hanging out with friends, sometimes the goal is the activity of sharing an experience itself.

It is in this sense that we can understand the reciprocal interactions between infants and primary caretakers as joint action. They are not directed towards bringing about an external goal, but infants still communicate and interact with their primary caretakers in coordinated ways and feel rewarded when these early attempts to reach out and engage in shared experiences meet their expectations. I believe this points us towards an understanding of a different, and important, kind of joint action.

## **5. Conclusion**

In this chapter, I have argued that early reciprocal interactions between adult caretakers and infants are a form of joint action. These interactions occur before infants are nine months old and occur during the stage of primary intersubjectivity. I defended this view by first presenting empirical evidence that suggests that infants engage in reciprocal, mutually rewarding interactions from birth. These interactions, unlike joint attention or shared intentionality, are dyadic. This means that they involve only the infant and the caretaker and do not refer to a third, external object towards which the infant and her caretaker co-orient. Actions or states that involve this third element or triangular shape are, by contrast, triadic.

I then discussed some recent philosophical accounts of joint action in young children. These accounts were formulated in response to a problem generated by the cognitive complexity required by standard accounts of joint action as well as the reality that young children and infants

do engage in joint action. Because young children have not yet developed the robust theory of mind required to formulate the complex, interlocking intentions required by accounts such as Bratman's, new, less cognitively demanding accounts of joint action seemed to be required. However, while these accounts do make some important progress, they unnecessarily retain the idea that joint action involves the pursuit of an external goal or object of interest outside the participating agents. That is, they unnecessarily retain the idea that joint action is inherently triadic.

Next, I argued that these recent philosophical accounts of joint action in young children cannot explain the early reciprocal interactions of the infant-adult dyad. This is in part because the accounts restrict themselves to analyzing the child's ability to engage in triadic forms of shared activity. Because triadic forms of engagement involve an ability to represent the other as a subject of experiences that may diverge from one's own, to engage in means-end reasoning, and to engage in socially recursive mind-reading processes, infants under nine months old are excluded from these analyses.

However, there is good reason to think that triadic forms of interaction do not exhaust the infant's ability to act jointly with others. I concluded by suggesting that there are dyadic forms of joint action that only require participants to be able to share an intersubjective state with the other while coordinating one's actions. Because even infants seem capable of meeting these conditions when they reciprocally imitate or engage in proto-conversations with an adult caretaker, this account potentially provides us with a way of understanding joint action that does not conform to the standard, triadic structure at every stage of life.

## References

- Bateson, Mary Catherine. 1975. "Mother-Infant Exchanges: The Epigenesis of Conversational Interaction." *Developmental Psycholinguistics and Communication Disorders* 263 (1): 101-13.
- Bateson, Mary Catherine. 1979. "The Epigenesis of Conversational Interaction: A Personal Account of Research Development." In *Before Speech: The Beginning of Human Communication*, edited by M. Bullowa, 63-77. London: Cambridge University Press.
- Batki, Anna, Simon Baron-Cohen, Sally Wheelwright, Jennifer Connellan, and Jag Ahluwalia. 2000. "Is There an Innate Gaze Module? Evidence from Human Neonates." *Infant Behavior and Development* 23 (2): 223-9.
- Bratman, Michael. 1999. *Faces of Intention*. Cambridge: Cambridge University Press.
- Brinck, Ingar, Vasudevi Reddy, and Dan Zahavi. 2017. "The Primacy of the "We"?" In *Embodiment, Enaction, and Culture: Investigating the Constitution of the Shared World*, edited by C. Durt, T. Fuchs and C. Tewes, 131-47. Cambridge: MIT Press.
- Bushnell, Ian, F. Z. Sai, and J. T. Mullin. 1989. "Neonatal Recognition of the Mother's Face." *British Journal of Developmental Psychology* 7 (1): 3-15.
- Butterfill, Stephen. 2012. "Joint Action and Development." *The Philosophical Quarterly* 62 (246): 23-47.
- Cohn, Jeffrey F., and Edward Z. Tronick. 1983. "Three-Month-Old Infants' Reaction to Simulated Maternal Depression." *Child Development* 54 (1): 185-93.
- Csibra, Gergely. 2010. "Recognising Communicative Intentions in Infancy." *Mind & Language* 25 (2): 141-68.

- Fantasia, Valentina, Gabriela Markova, Alessandra Fasulo, Alan Costall, and Vasudevi Reddy. 2016. "Not Just Being Lifted: Infants are Sensitive to Delay During a Pick-Up Routine." *Frontiers in Psychology* 6 (article 2065): 1-8. <https://doi.org/10.3389/fpsyg.2015.02065>.
- Fantz, Robert L. 1963. "Pattern Vision in Newborn Infants." *Science* 140 (3564): 296-7.
- Farroni, Teresa, Gergely Csibra, Francesca Simion, and Mark H. Johnson. 2002. "Eye Contact Detection in Humans from Birth." *Proceedings of the National Academy of Sciences of the United States of America* 99 (14): 9602-5.
- Farroni, Teresa, Enrica Menon, and Mark H. Johnson. 2006. "Factors Influencing Newborns' Preference for Faces with Eye Contact." *Journal of Experimental Child Psychology* 95 (4): 298-308.
- Farroni, Teresa, Mark H. Johnson, and Gergely Csibra. 2004. "Mechanisms of Eye Gaze Perception during Infancy." *Journal of Cognitive Neuroscience* 16 (8): 1320-6.
- Meltzoff, Andrew N., and M. Keith Moore. 1977. "Imitation of Facial and Manual Gestures by Human Neonates." *Science* 198 (4312): 75-78.
- Meltzoff, Andrew N., and M. Keith Moore. 1983. "Newborn Infants Imitate Adult Facial Gestures." *Child Development* 54 (3): 702-9.
- Moll, Henrike, Ellyn Pueschel, Qianhui Ni, and Alexandra Little. 2021. "Sharing Experiences in Infancy: From Primary Intersubjectivity to Shared Intentionality." *Frontiers in Psychology* 12 (667679): 1-13. <https://doi.org/10.3389/fpsyg.2021.667679>.
- Moran, Richard. 2018. *The Exchange of Words: Speech, Testimony and Intersubjectivity*. Oxford: Oxford University Press.
- Nagy, Emese. 2011. "The Newborn Infant: A Missing Stage in Developmental Psychology." *Infant and Child Development* 20 (1): 3-19.

- Nagy, Emese, and Peter Molnar. 2004. "Homo Imitans or Homo Provocans? Human Imprinting Model of Neonatal Imitation." *Infant Behavior & Development* 27 (1): 54-63.
- Pacherie, Elisabeth. 2013. "Intentional Joint Agency: Shared Intention Lite." *Synthese* 190 (10): 1817-39.
- Reddy, Vasudevi. 2008. *How Infants Know Minds*. Cambridge: Harvard University Press.
- Reddy, Vasudevi. 2011. "A Gaze at Grips With Me." In *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience*, edited by A. Seemann, 137-57. Cambridge: MIT Press.
- Reddy, Vasudevi, Gabriela Markova, and Sebastian Wallot. 2013. "Anticipatory Adjustments to Being Picked Up in Infancy." *PLoS One* 8 (6).  
<https://doi.org/10.1371/journal.pone.0065289>.
- Rochat, Philippe. 2001. *The Infant's World*. Cambridge: Harvard University Press.
- Rochat, Philippe, Claudia Passos-Ferreira, and Pedro Salem. 2009. "Three Levels of Intersubjectivity in Early Development." In *Enacting Intersubjectivity. Paving the Way for a Dialogue Between Cognitive Science, Social Cognition and Neuroscience*, edited by Antonella Carassa, Francesca Morganti, and Giuseppe Riva, vol. 10, 173-90. Lugano: Universita Svizzera Italiana.
- Roth, Abraham Sesshu. 2004. "Shared Agency and Contralateral Commitments." *Philosophical Review* 113 (3): 359-410.
- Roth, Abraham Sesshu. 2014. "Prediction, Authority, and Entitlement in Shared Activity." *Nous* 48 (4): 626-52.

- Sander, Louis. 2007. *Living Systems, Evolving Consciousness, and the Emerging Person: A Selection of papers from the Life Work of Louis Sander*. New York and London: Routledge.
- Schmid, Hans Bernhard. 2014. "Plural Self-Awareness." *Phenomenology and the Cognitive Sciences* 13 (1): 7-24.
- Searle, John. 1983. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Stern, Daniel. 1971. "A Micro-Analysis of Mother-Infant Interaction: Behaviors Regulating Social Contact Between a Mother and Her Three-and-a-Half-Month-Old twins." *Journal of the American Academy of Child Psychiatry* 10 (3): 501-17.
- Stern, Daniel. 1974. "Mother and Infant at Play: The Dyadic Interaction Involving Facial, Vocal and Gaze Behaviours." In *The Effect of the Infant on its Caregiver*, by M. Lewis and L. A. Rosenblum, 187-213. New York: Wiley.
- Stern, Daniel. 1985. *The Interpersonal World of the Infant: A View from Psychoanalysis and Development Psychology*. New York: Basic Books.
- Stern, Daniel. 1999. "Vitality Contours: The Temporal Contour of Feelings as a Basic Unit for constructing the infant's social experience." In *Early social cognition: Understanding others in the first months of life*, by P. Ed. Rochat, 67-90. NJ: Erlbaum Associates Publishers.
- Stern, Daniel. 2008. "The Clinical Relevance of Infancy: A Progress Report." *Infant Mental Health Journal* 29 (3): 177-88.
- Story, Daniel. 2021. "Joint Action Without Robust Theory of Mind." *Synthese* 198:5009-5026.

- Tollefsen, Deborah. 2005. "Let's Pretend! Children and Joint Action." *Philosophy of the Social Sciences* 35 (1): 75-97.
- Tomasello, Michael. 1995. "Joint Attention as Social Cognition." In *Joint Attention: Its Origin and Role in Development*, edited by Chris Moore and Philip J. Dunham, 103-30. New Jersey: Lawrence Erlbaum Associates, Inc.
- Tomasello, Michael. 1999. *The Cultural Origins of Human Cognition*. Cambridge: Harvard University Press.
- Tomasello, Michael. 2018. "Great Apes and Human Development: A Personal History." *Child Development Perspectives* 12 (3): 189-93.
- Tomasello, Michael. 2019. *Becoming Human: A Theory of Ontogeny*. Cambridge: Belknap Press.
- Tomasello, Michael, and Malinda Carpenter. 2007. "Shared Intentionality." *Developmental Science* 10 (1): 121-5.
- Tomasello, Michael, Malinda Carpenter, Josep Call, Tanya Behne, and Henrike Moll. 2005. "Understanding and Sharing Intentions: the Origins of Cultural Cognition." *Behavioral and Brain Sciences* 28 (5): 675-91.
- Tomasello, Michael, and Henrike Moll. 2010. "The Gap is Social: Human Shared Intentionality and Culture." In *Mind the Gap*, edited by P.M. Kappeler and J. B. Silk, 331-49. Berlin: Springer-Verlag.
- Trauble, Birgit, and Sabina Pauen. 2011. "Cause or Effect: What Matters? How 12-month-old Infants Learn to Categorize Artifacts." *British Journal of Developmental Psychology* 29 (pt. 3): 357-74.

- Trevarthen, Colwyn. 1979. "Communication and Cooperation in Early Infancy. A Description of Primary Intersubjectivity." In *Before Speech: The Beginning of Human Communication*, edited by M. Bullowa, 321-47. London: Cambridge University Press.
- Trevarthen, Colwyn. 2011a. "The Generation of Human Meaning: How Shared Experience Grows in Infancy." In *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience*, edited by A. Seemann, 73-113. Cambridge: MIT Press.
- Trevarthen, Colwyn. 2011b. "What Is It Like to Be a Person Who Knows Nothing? Defining the Active Intersubjective Mind of a Newborn Human Being." *Infant and Child Development* 20 (1): 119-35.
- Trevarthen, Colwyn, and Ken Aitken. 2001. "Infant Intersubjectivity: Research, Theory, and Clinical Applications." *Journal of Child Psychology and Psychiatry* 42 (1): 3-48.
- Tronick, Edward Z., and Jeffrey F. Cohn. 1989. "Infant-Mother Face-to-Face Interaction: Age and Gender Differences in Coordination and the Occurrence of Miscoordination." *Child Development* 60 (1): 85-92.
- Vincini, Stefano, Yuna Jhang, Eugene H. Buder, and Shaun Gallagher. 2017. "Neonatal Imitation: Theory, Experimental Design, and Significance for the Field of Social Cognition." *Frontiers in Psychology* 8 (article 1323): 1-16.  
<https://psycnet.apa.org/doi/10.3389/fpsyg.2017.01323>.
- Wellman, Henry M. 2002. "Understanding the Psychological World: Developing a Theory of Mind." In *Blackwell Handbook of Childhood Cognitive Development*, edited by Usha Goswami, 167-87. New Jersey: Blackwell Publishing.



## Conclusion

In this dissertation, I have analyzed atypical cases of joint action found in both direct and indirect acts of telling as well as the reciprocal interactions that occur between infants and their adult caretakers at the earliest stages of life. These cases are atypical, I have argued, because they resist analyses of joint action that treat the phenomenon as something that occurs in primarily cooperative contexts, in which individual participants share a goal that they endeavor to pursue by forming shared intentions and being mutually responsive to the plans and actions of their co-participants. This is because speech acts like telling and insinuation and the early reciprocal interactions that take place between infants and adult caretakers do not seem to conform to this basic structure. Telling and insinuation often occur in highly antagonistic contexts in which the speaker and interlocutor have very few goals in common outside of speaking English or communicating. Early interactions between infants and caretakers lack the shared intentions and mutual responsiveness that many accounts of joint action emphasize.

These atypical cases of joint action suggest that the tendency in the philosophical literature of focusing on highly cooperative and structured examples can misrepresent the phenomenon. It can lead us to believe that joint action is synonymous with highly cooperative interactions between two agents who are mutually responsive to the plans and goals of their co-participants so that they do not infringe on one another's freedom. But, as I have tried to show throughout this dissertation, this is a mistake, one which can undermine our understanding of joint action by limiting us to just one aspect of the phenomenon. Attaining a more comprehensive understanding of joint action will require further analysis into these undertheorized cases to determine what, if anything, they have in common with other forms of joint action, and whether they might inform our understanding of even the standard cases.