Check for updates

RESEARCH ARTICLE

REVISED Skilled we-intentionality: Situating joint action in the living environment

[version 2; peer review: 2 approved]

Julian Kiverstein [1], Erik Rietveld [1-3]

[1]Psychiatry, Academic Medical Center of the University of Amsterdam, Amsterdam, 1105AZ, The Netherlands
[2]Philosophy, University of Twente, Enschede, The Netherlands
[3]Institute of Logic, Language and Computation, Humanities, University of Amsterdam, Amsterdam, The Netherlands

v2 **First published:** 21 May 2021, **1**:54
https://doi.org/10.12688/openreseurope.13411.1
Latest published: 27 Sep 2021, **1**:54
https://doi.org/10.12688/openreseurope.13411.2

Abstract

There is a difference between the activities of two or more individuals that are performed jointly such as playing music in a band or dancing as a couple, and performing these same activities alone. This difference is sometimes captured by appealing to shared or joint intentions that allow individuals to coordinate what they do over space and time. In what follows we will use the terminology of we-intentionality to refer to what individuals do when they engage in group ways of thinking, feeling and acting. Our aim in this paper is to argue that we-intentionality is best understood in relation to a shared living environment in which acting individuals are situated. By the "living environment" we mean to refer to places and everyday situations in which humans act. These places and situations are simultaneously social, cultural, material and natural. We will use the term "affordance" to refer to the possibilities for action the living environment furnishes. Affordances form and are maintained over time through the activities people repeatedly engage in the living environment. We will show how we-intentionality is best understood in relation to the affordances of the living environmentand by taking into account the skills people have to engage with these affordances. For this reason we coin the term 'skilled we-intentionality' to characterize the intentionality characteristic of group ways of acting, feeling and thinking.

Keywords

Joint action; we intentionality; skilled intentionality; affordances; dynamical systems theory; embodied cognition; philosophy of action

Open Peer Review

Approval Status ✓ ✓

	1	2
version 2 (revision) 27 Sep 2021	✓ view	✓ view
version 1 21 May 2021	? view	? view

1. **Glenda Satne**, University of Wollongong, Wollongong, Australia

2. **Anthony Chemero**, University of Cincinnati, Cincinnati, USA

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the Philosophy, Ethics and Religion gateway.



This article is included in the Horizon 2020 gateway.



This article is included in the Psychology gateway.



This article is included in the European Research Council (ERC) gateway.

Corresponding author: Julian Kiverstein (j.d.kiverstein@amsterdamumc.nl)

Author roles: Kiverstein J: Conceptualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Rietveld E**: Conceptualization, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No [679190]), (Skilled Intentionality for 'Higher' Embodied Cognition: Joining forces with a field of affordances in flux [AFFORDS-HIGHER]), and by an Amsterdam Brain and Cognition Research Grant awarded to Damiaan Denys and Erik Rietveld.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2021 Kiverstein J and Rietveld E. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Kiverstein J and Rietveld E. **Skilled we-intentionality: Situating joint action in the living environment [version 2; peer review: 2 approved]** Open Research Europe 2021, **1**:54 https://doi.org/10.12688/openreseurope.13411.2

First published: 21 May 2021, **1**:54 https://doi.org/10.12688/openreseurope.13411.1

REVISED **Amendments from Version 1**

In responding to reviewer 1 we have aimed at clarifying the explanatory ambitions of our paper. Our ambition was to provide a framework for understanding we-intentionality in its entirety, and not only the more modest goal of providing a philosophical account of action coordination in dyads or groups. We have foregrounded the role that situated normativity plays in our account, and show how this allows us to handle the counterexamples proposed by reviewer 1. Second, we have added new material that addresses the different projects in the literature on we-intentionality reviewer 1 usefully distinguished.

In response to reviewer 2, we have better emphasised the places in our paper where we see ourselves as making important new contributions to the literature. There are three main contributions of our paper. First we show how to understand interpersonal synergies in relation to a field of relevant affordances. Second, we introduce the concept of large-scale affordances from our earlier work to make sense of how agents can coordinate their actions over long time-scales without appealing to the notion of planned coordination. Finally, we show how the concept of affordance from ecological psychology is much broader in scope than has hitherto been recognised, and may generalise to play a role in the explanation of we-intentionality, a form of intentionality that has been argued to be uniquely human by Michael Tomasello among others. Our paper thus makes an original contribution to the literature on the scaling-up problem for embodied cognitive science, showing how such explanations can be brought to bear to address an allegedly unique form of human cognition that is typically thought to call for higher-order reasoning about mental states.

Any further responses from the reviewers can be found at the end of the article

Plain language summary

How could the environment be structured to allow people to act as groups?

People act together all the time. Think of a couple painting their house together or carrying a heavy object like a piano up a steep flight of stairs in an Amsterdam apartment. These are the kind of activities we aim to understand in this philosophical article. In philosophy researchers have written a good deal about how it is possible for two or more agents to think or act as members of a group. To carry the piano up the stairs, for instance, people must avoid getting in each other's way or pulling in different directions.

Our research group investigates how people structure the environment in ways the support their actions. In this article our aim is to show how the environment also has a crucial role to play when people do things together. The stairs and the weight of the piano for instance tightly constrain how removal men move their bodies as they heave the piano up the stairs. The main contribution of this article is to describe the much overlooked reliance agents have on the structure in the environment when they act together.

Introduction

People take part in collaborative and joint activities. They help each other to move house. They play games against each other and make music together. They act in solidarity with each other to protest against social injustice. In each of these examples people are doing things together; they are acting jointly. Joint actions are different from co-actions in which multiple individuals each do their own thing. Philosophers have accounted for this difference by appealing to a special class of intentional states referred to variously as we-intentions, shared-intentions, or joint-intentions (Bratman, 1993; Butterfill, 2012; Gilbert, 1989; Gilbert, 1996; Pacherie, 2007; Sellars, 1963; Searle, 1990; Tuomela & Miller, 1988; Tuomela, 2005). In what follows we will use the terminology of 'we-intentionality' to refer to what individuals do when they think, act, or feel as a member of a group. We take thinking, acting and feeling to be interrelated dimensions of skilled action (Dewey, 1934; cf. Rietveld *et al.*, 2018). We-intentionality, as we will use this term, is a characteristic of skilled actions performed collaboratively and collectively with other agents. Our aim in this paper is to argue that we-intentionality is best understood in relation to a shared living environment in which acting individuals are situated. By the 'living environment' we mean to refer to places and everyday situations in which humans act. These places and situations are simultaneously social, cultural, material, and natural. We will use the term 'affordance' to refer to the possibilities for action the living environment furnishes (Gibson, 1979/2014). Affordances form and are maintained over time through the activities people repeatedly perform in the living environment (Kiverstein *et al.*, 2019; Rietveld & Kiverstein, 2014; Van Dijk & Rietveld, 2017). We-intentionality is, we claim, best understood in relation to the affordances of the living environment and by taking into account the skills people have to engage with these affordances.

Philosophers and cognitive psychologists have typically approached we-intentionality by asking what it takes for individuals to perform an action together and to share an intention or goal to do so. Actions are the doings of agents, they are not passively undergone by agents. It is the person's intention that is standardly taken to distinguish an action of an agent from an accident for which the agent is excused. Some of the actions that agents can control are joint actions two or more agents perform together. Philosophers have suggested that what is distinctive of these joint actions is that they are caused by intentions and goals that are shared among the individuals participating in the action. The philosophical debate has been concerned with the conditions under which two or more individuals can be said to share an intention or goal.

However, the focus on this question has led philosophers to neglect the contribution of the living environment to we-intentionality. Infants grow-up in an environment structured by caregivers to facilitate the infant's development of skills for participating in the daily routines of their community. They learn about what can be done with the things around

them, in part through interaction with other people in structured social settings (Reed, 1996, ch.9 & 10). The ecological psychologist Edward Reed described how other people create what he called a "field of promoted action" for infants (Reed, 1991)[1]. What is promoted to the infant are everyday action possibilities that allow the infant to understand, develop skills, and eventually take part in the practices of their community. The education of the child's attentions to those affordances results in joint attention (also see Segundo-Ortin & Satne, forthcoming). Part of the experience the child and caregiver undergo is that they are together attending to a possibility for action. What the child learns in the process is that affordances are not only for individuals but allow for the pursuit of joint activities such as cooking, tidying, playing, or making things together.

Drawing upon these important insights from the research tradition of ecological psychology, we will provide an account of we-intentionality in terms of a skilled agent's selective responsiveness to relevant affordances, or what we propose to call *skilled we-intentionality*. When people engage in activities together, we will argue, it is the shared relevant affordances of the living environment that explain how they manage to coordinate with each other. In acting together, agents dynamically couple to each other in ways that are constrained and scaffolded by the affordances of shared relevance to them.

Our paper consists of four sections. In section one, we introduce the central idea of our paper that we-intentionality should be understood in relation to the affordances of the living environment. In section two we introduce the concept of skilled intentionality as selective responsiveness to multiple relevant affordances. In section three we show how, when people act jointly, they are able to make their actions coordinate with each other over space and time because they are responsive to affordances of shared relevance. In section four we explain what it is for affordances to be of shared relevance to multiple individuals. We conclude that group modes of thinking and acting are made possible by affordances of the living environment that are of shared relevance to multiple individuals.

The role of affordances in action coordination

We-intentionality is standardly taken to be a property of the mental states that account for how individuals can think, act and feel as members of a group or communal practice[2]. We will use the term to refer to what agents do when they skilfully engage in and contribute to group and communal activities such as playing music in a band, taking part in a team sport, working together on a collaborative project, or acting in agreement with a local custom. These activities require skill and sensitivity on the part of the agent. The agent must sensitively adapt and adjust what they do when they take part in group or communal activities so that it fits with the demands of the particular circumstances in which they are acting. They must be sensitive to whether their performance is unsatisfactory or adequate, appropriate or inappropriate relative to what others in the group expect of them[3].

The term 'we-intentionality' can be applied at a variety of scales from joint action at the scale of the dyad, to collective patterns of skilled action at the scale of groups and communities. The conceptual framework for understanding we-intentionality we propose is intended to generalise across these scales. This is reflected in the examples we have chosen, some of which are examples of dyads performing activities together. Other examples concern communal activities in which individuals regulate their actions so as to make them conform with the patterned practices of their communities. Our aims in this paper are thus general. We propose a conceptual framework – the Skilled Intentionality Framework (Rietveld *et al.*, 2018) – for making sense of we-intentionality. The task of applying this framework to the rich variety of phenomena philosophers have

[1] The concept of the field of promoted action is closely related to what the developmental psychologist Lev Vygotsky referred to as the "zone of proximal development" (Vygotsky, 1978) and Jerome Bruner called "scaffolding" (Bruner, 1983; Rogoff, 1990). Reed's idea was that caregivers promote actions to the child that are beyond what the child can already do. The infant is encouraged to act in ways that it cannot yet quite manage so that the child becomes familiar with the meaning and value of the affordances of her surroundings. Think of how children are introduced to story books way before they are able to recognise that the words written in these books mean something. The infant recognises that something meaningful is going on in the book before she is able to understand what it is.

[2] The comparative psychologist Michael Tomasello makes a distinction between two forms of we-intentionality he calls "joint" and "collective intentionality" (Tomasello, 2014). Joint intentionality attaches to small-scale collaborative actions and shared attention between individuals in the here and now. What makes the interaction joint is that it is characterised by a second-person engagement: an I-you relation that the individuals recognise they are taking part in. Tomasello uses the term "collective intentionality" for larger-scale forms of collaboration that take the form of conventions, norm-governed practices and institutions. While early human collaborations were second-personal modern humans added to this form of social interaction group minded ways of thinking and acting. Collective intentionality made possible agent-neutral, objective ways of thinking and evaluating that rely upon common ground – what everyone in the group counts on everyone else knowing.

[3] We have elsewhere referred to this sensitivity as 'situated normativity' – the ability of skilled individuals to distinguish better from worse, adequate from inadequate, appropriate from inappropriate, or correct from incorrect in the context of a particular situation (Rietveld, 2008). This ability has been shaped by a history of taking part in sociomaterial practices. We describe this ability as 'situated' because the embodied sensitivity the individual develops by learning to act skilfully is a sensitivity they have cultivated through their participation in sociomaterial practices. Such skills can also be exercised in situations that call for reflection, or what we have called reflective situated normativity (Van Den Herik & Rietveld, 2021). Reflective situated normativity is a form of normativity that relies on linguistic practices and involve explicit articulation of rules. If I tell my children: "No phones at the dinner table!", this is an example of reflective situated normativity. The articulated rule attracts attention to, or *puts on display* (cf. Noë, 2015) a pattern of behaviour as inappropriate. It can therefore be characterised as reflective.

discussed under the heading of we-intentionality goes beyond what we can achieve in this paper.

We propose to use the term 'we-intentionality' in this *scale-able*, and therefore slightly non-standard way, in part on developmental grounds. Shared ways of thinking, feeling and acting can already be observed in dyadic engagement between infants and their caregivers in family life (Hobson, 2002; Reddy, 2018; Rochat & Robbins, 2017; Rakoczy, 2017; Vincini & Gallagher, 2021; Satne & Sallice, 2020; Segundo-Ortin & Satne, forthcoming). The collective acceptance of the rules and norms of communal social practices and cultural institutions and conventions may build upon the ability to coordinate and attune to the actions and reactions of others that emerge in early development. Think for instance of emotional tuning to other's approval or disapproval that underpin a child's developing a sense of the customs and norms of her community (Haugeland, 1990; Satne, 2014). Our adult abilities to make sense of complex social institutions such as governments, universities, and corporations builds upon and grows out of abilities to blindly, but nevertheless sensitively, follow shared norms that emerge very early in development.

As noted in our introduction, many human skills are mastered by children not by themselves but in social interaction with their caregivers. Many of the problems the infant runs into in skill development are thrust upon the infant by its caregivers (Reed & Bril, 1996). The caregiver calls the attention of the infant to particular affordances provided by their living environment. Think for instance of the motor problems the infant encounters in trying to handle eating implements such as spoons or chopsticks to feed themselves while eating with their family. The motor problems in this case originate in the skills the child needs to develop for engaging in a communal activity of eating together. These skills make possible what we are calling we-intentionality defined as what agents do when they engage in and contribute to a group or communal activity. The motor problems the infant encounters have a particular form because of the multiple affordances it needs to coordinate with to take part in a family meal. In the process of being introduced and promoted to the child by its caregivers, these affordances become joint affordances that will eventually enable the child to take part in this shared activity. This is a simple illustration of the central claim of our paper that we-intentionality should be understood in relation to the affordances of the living environment.

The living environment has been largely absent from the philosophical discussion of we-intentionality[4]. The debate surrounding the nature of we-intentionality has been mostly focused on the question of whether we-intentions are reducible to complex aggregates of interlocking attitudes of individual agents. Searle (1990) for example, argued against such a reduction of we-intentions to the intentions of individuals. For Searle we-intentions are distinguished by a psychological mode "that must make reference to collectives" (1990: p.408). Bratman (1993) by contrast denies that we-intentions form a special, psychologically primitive class of intentions. He has provided a reductive account of what he calls "shared-intentions" in terms of the interlocking intentions and attitudes of the participants in a joint action. It is the interrelatedness of the individual's intentions and the complementarity of their plans that Bratman takes to be explanatory of joint actions. We-intentions can be distinguished from the intentions of individuals because they require there to be certain relations between the intentions and attitudes of individuals and for the individuals to know these relations obtain.

Bratman's account is tailor-made for cases of conscious and deliberate action in which agents act in pursuit of some shared plan formed prior to the execution of a joint action. Plans are conceived by Bratman as supporting the coordination of various activities over long periods of time. The formation of such a long-term plan requires practical reasoning to work out the course of action that is conducive to bringing about this plan. It may well require the formation of a hierarchy of sub-plans and intentions for bringing about each of these sub-plans. The account of interlocking intentions Bratman provides is perhaps well-suited to account for the negotiations that might unfold between people when they plan to perform an action together (e.g., Bratman, 2014). His account does not aim to explain how multiple agents succeed in putting the plans they have made into effect, coordinating what they are doing when they act together in a particular situation.

Searle's account of we-intentions also has nothing to say about how the actions of different individuals must fit together in order for them to successfully perform a joint action. Searle seem to think the answer to this question can be given by appeal to biological background capacities that are not themselves intentional phenomena (Searle, 1995, ch.6). However, this division of labour between philosophy and science will not do. Searle's distinction between we-intentions and non-intentional background capacities introduces a problematic divide within the mind between 'lower-level' mechanical coordinative structures and 'higher-level' we-intentions and goals. The details of how a we-intention is implemented by multiple individuals as they act together in a particular situation are treated by Searle as a matter of the causal capacities of the brain that enables, we-intentional states (Searle, 1995: p.129). We question whether this causal separation of 'lower-level' coordinative processes from 'higher-level' we-intentions can be made in practice (cf. Pacherie, 2007).

Butterfill (2012) correctly rejects Searle's separation of we-intentions from non-intentional background capacities when he describes how coordinative structures form across

[4] Some notable exceptions are Abramova & Slors (2015); Fusaroli & Tylén (2012); Knoblich *et al.* (2011) and Gallagher & Ransom (2016). Knoblich *et al.* (2011) and Abramova & Slors (2015) both foreground the role of what they call "joint affordances" in action coordination. We briefly discuss their proposal later in this section. Fusaroli & Tylén argue that semiotic systems provide shared affordances that constrain and stabilise linguistic and non-linguistic coordination. Gallagher and Ransom usefully deploy the Material Engagement Theory of Lambros Malafouris (2013) to highlight the role that artefacts play in action coordination.

two or more bodies through the coupling of perception and action (cf. Richardson *et al.*, 2009; Shockley *et al.*, 2009; Tollefsen & Dale, 2012). He argues for a class of plural activities that require only a shared goal or outcome. A group of ants for instance can work together to kill a large insect. The ant's individual behaviours are organised to bring about the death of the insect but they have this effect only through their coordination. Butterfill introduces the concept of a shared goal to explain how the actions of the individual ants could be coordinated. Their actions can be directed at one and the same goal as a consequence of what Butterfill calls 'emergent coordination'. Coordination is the outcome of perception-action couplings between agents and does not require the agents to form plans or to reason about each other's intentions. Multiple agents can begin to act together because the actions of each of them are responsive to the same environmental cues and motor routines, and each agent makes more or less the same contribution to bringing about a shared goal.

As an example of emergent coordination consider the following experiment in which participants were given the task of lifting planks from a moving conveyor belt (Richardson *et al.*, 2007; Richardson *et al.*, 2010). Participants were told that they were only allowed to lift the planks by their ends, not from the centre which meant that some of the planks were too long to be lifted alone, and had to be carried by two participants working together. The experimenters found that people switched from lifting the planks alone, to working as a dyad based on the ratio of their arm span to the plank length (Richardson *et al.*, 2007: 849). As the planks got closer to an individual's maximum arm span, so they spontaneously switched from moving the plank alone to moving the plank together with the other agent. The presence of the other person in this case induces emergent coordination. In the presence of the other person the plank becomes "jointly liftable" or "liftable by us" (Knoblich *et al.*, 2011; cf. Abramova & Slors, 2015: p.528–9). The possibility to lift the long plank is an affordance for the dyad comprised of the two co-actors when they work together.

There are two important points we wish to highlight from this example of emergent coordination. Firstly, what Butterfill calls the 'shared goals' of the individuals should be understood in relation to the affordances of the plank (Abramova & Slors, 2015). Recall that one of the questions that has occupied philosophers working on we-intentionality is how the intentions of distinct individuals can be coordinated over space and time so that the individuals can take part in a joint action. Reductive accounts like that of Bratman have supposed that we-intentions must somehow represent who will do what and when and that such a representation must be common knowledge among the agents taking part in the joint action. However, it is less clear that this kind of complex "mental gymnastics" (Chemero, 2009) is necessary once we think of intentions in relations to the affordances of the environment. We expand on this point in the next section.

Secondly, when the two individuals work together to the lift the plank they can be thought of as forming an interpersonal synergy. The notion of synergy here is that of a constraint that when applied to a system causes the elements of a system to form units that work together. Chemero (2015) gives as an example of a synergy laser in which "large amounts of energy constrain photons so that they coordinate their behaviour with another over long spatial and temporal scales" (p.143). The affordance of liftability can be thought of as a constraint that enables the two individuals to temporarily form an interpersonal synergy and act together as a single unit. We return to this point in section three. Once we keep these two points in mind, the problem of how two or more individuals can act together looks very different than what we saw above in for example, the accounts of Searle and Bratman. Coordination is something that unfolds out in the open between them, instead of being something that can only be achieved through complex reasoning about each other's mental states.

The Skilled Intentionality Framework: explaining goals and intentions

Our aim in this section is to show how the work that shared goals and intentions are asked to do in philosophical accounts of joint action can instead be performed by affordances of shared relevance. We provide an analysis of intentions and goals in terms of what we will call 'skilled-intentionality' (Bruineberg & Rietveld, 2014; Kiverstein & Rietveld, 2021; Rietveld *et al.*, 2018; Van Dijk & Rietveld, 2017). The concept of skilled intentionality is a generalisation and elaboration of what Merleau-Ponty (1945/2002) called 'motor intentionality' to apply to skilled expertise in general. Merleau-Ponty talks about the skilled typist for instance as having knowledge in her hands. Based on her motor skills she experiences the keyboard as inviting certain movements of her fingers as she types (1945/2002: 166). Similarly, the goalkeeper's football skills allow them to time their dive just right, so as to intercept a ball heading towards their goal.

Merleau-Ponty describes the actions of the skilled agent as being drawn forth from them by the world. The football field is for instance "pervaded by lines of force (the 'yard' lines; those which demarcate the 'penalty area') … which call for a certain mode of action." (Merleau-Ponty, 1943/1962: 168) We've used the different terminology of the 'field of relevant affordances' to capture the same idea of skilled actions being solicited or invited by the world[5]. 'Relevant affordances' are the possibilities for action offered by the environment that 'stand-out' in the sense of being marked out as salient or significant from the other affordances the living environment provides (cf. Withagen *et al.*, 2012). If I am currently hungry for instance, the apple that is sitting by my computer will appear more alluring than when I have just eaten.

We will use the term the 'landscape of affordances' to refer to the totality of affordances that can be found in the ecological

[5] Our aim is to provide "bridging concepts" that allow us to connect phenomenology to empirical research in ecological dynamical cognitive science (Chemero, 2009; Kaufer & Chemero, 2015).

niche of a given form of life. The landscape of affordances is exceptionally rich in terms of the possible actions it offers. Yet individuals are typically ready to respond selectively to affordances in ways that are appropriate to the context, and that reflect their own internal states that arise in relation to their situation. How is it possible that individuals are only responsive to relevant affordances in each particular situation? The very same affordances can stand out as relevant to an individual in one situation, while in a different situation the same affordance does not move them at all. What accounts for this difference?

We do not think appealing to the agent's goals and intentions helps to answer this question (Van Dijk & Rietveld, 2020). It just invites the question: why does the agent have the goals and intentions that they do? We thus agree with Lucy Suchman when she writes:

> To characterise purposeful action as in accord with plans and goals is just to say again that it is purposeful and that *somehow*, in a way not addressed by the characterisation itself, we constrain and direct our actions…How we do that is the outstanding problem. Plans and goals do not provide the solution for that problem, they simply restate it. (Suchman, 1987: 47–8, cited by Heft, 2001: 311)

Rather than presupposing that the agent has pre-existing plans, intentions and goals, we seek an account of goals and intentions in terms of the self-organising dynamics of the agent situated in a landscape of affordances. A system is self-organising if it exhibits ordered and regular patterns of macroscopic behaviour as a result of the endogenously generated interactions among its component parts, without any centralised control (Bruineberg & Rietveld, 2014; Freeman, 2000; Kelso, 1995; Ladyman & Wiesner, 2020; Strogatz, 1994; Tschacher & Haken, 2007; Tsuda, 2001). The self-organising system that we take to be explanatory of skilled intentionality is the whole animal-environment system.

The animal's responsiveness to affordances can be described in terms of an attractor landscape[6] with a particular form determined by the animal's learning history and the skills and abilities it has developed over the course of this history. Large scale patterns of activity spontaneously form in the brain and the rest of the body as a consequence of a history of engagement in a skilled activity. Certain patterns of activity are selected and preserved because they have proven to be practically useful to the organism in the past for coordinating with affordances (Reed, 1996). The organism's moves towards a basin of attraction as it prepares to act. Each basin of attraction corresponds to an affordance-related state of action readiness.

The affordances that stand out as relevant to an agent are those that are sensed as affectively significant because they contribute in some way to improving its grip. We borrow the notion of grip from Merleau-Ponty (1945/2002) to refer to the way the skilled individual is ready to move so as to improve its relation to the environment. Think of turning the dial on a microscope to get a better view of a biological specimen as an example of tending towards an optimal grip. So long as the specimen looks blurry through the microscope one feels one is not quite seeing it right. Based on this feeling one turns the dial until the image one is seeing through the lens is improved (i.e., sharper, and better defined). To be a skilled agent is to be able to maintain a good grip on the multiple relevant affordances required for the performance of a skilled action. The strikers in a football team will be sensitive to the events that are playing out on the football pitch because of their years of practice in the game, ready to take advantage of opportunities to strike on goal when they arise. Any change in the game or in the body can be sensed as disattunement, a deviation away from good grip, which, if all goes well, the striker will be affectively moved to reduce thereby achieving a better grip.

The states of bodily action readiness that relevant affordances elicit are inherently affective[7]. The agent in acting skilfully so as to tend towards an optimal grip modifies his or her relation to the world in a way that is in line with what matters to them[8]. If all goes well the animal will not come to rest in any basin of attraction but will be ready to switch to another attractor if a course of action does not proceed as anticipated (Bruineberg & Rietveld, 2019; Dreyfus, 2007; Freeman, 1987; Freeman, 2000). In this way the skilled agent doesn't just tend towards a single attractor basin at any given moment but is simultaneously influenced by multiple attractors. They are ready to transit between multiple attractors and can therefore be ready to act on multiple affordances simultaneously.

Now that we have the concept of skilled intentionality in place, let us consider again the debate between Searle, Bratman and others we discussed in the previous section concerning the reducibility or otherwise of we-intentionality to the interlocking mental states of individuals. What is in question in this debate is the possibility of explaining we-intentionality in terms that do not presuppose the sharing of mental states. In the Skilled Intentionality Framework, it is the shared affordances of the living environment that are appealed to in making sense of the coordination of the activities of multiple individuals. The sharing of affordances come in in two ways in our account. First, the landscape of affordances is defined in relation to a *shared* form of life. The affordances that the landscape

[6] The concept of an attractor landscape is drawn from dynamical systems theory (see e.g. Beer, 2000) in which the behaviour of a system over time is mapped onto an abstract geometric space. An attractor refers to a location within this space such that any trajectory (a temporal sequence of states) that passes nearby this point will be sucked towards it. A basin of attraction is the area of the state space within which the attractor exerts its influence. Adjoining basins of attractions form what we have referred to above as an "attractor landscape".

[7] Here we are drawing upon the analysis of emotion found in the research of the emotion psychologist Nico Frijda. See (Frijda, 1986; Frijda, 2007).

[8] In line with current ideas in computational neuroscience (Friston, 2010; Friston, 2011) we've argued that the tendency towards optimal grip is a consequence of the individual agent regulating their interactions with the environment in such a way as to continuously minimise "free energy" over time (Bruineberg & Rietveld, 2014; Bruineberg & Rietveld, 2019). In this paper we do not have the space to relate our argument to theoretical neurobiology, and leave this as a task for another occasion).

makes available have formed through patterns of practice that are constituted by the coordinated behaviour of multiple individuals. Individuals and collectives of the given form of life are situated in the same landscape of affordances; they share the same landscape of affordances.

Second, interpersonal synergies form on the basis of affordances of *shared* relevance, as we will see in more detail in the next section. The abilities that are pooled are abilities for distinguishing better from worse, adequate from inadequate, appropriate from inappropriate, or correct from incorrect in the context of a particular situation in a given sociomaterial practice[9]. We use the term 'situated normativity' to refer to this dimension of skilled action (Rietveld, 2008; van den Herik & Rietveld, 2021). It is based on the pooling of abilities for making these kinds of normative distinctions that individuals are able to tend towards an improved grip on a situation in acting together as a group. Crucially however, the account we provide of *skilled we-intentionality* does not make appeal to a *sui generis* or presupposed type of shared or collective mental state. This is because all cases of skilled intentionality, as we show in the next section, require coordinating with other members of a practice. Skilled intentionality is always already taking into account the established sociomaterial practices because affordances owe their existence in part to a shared social, cultural, material and natural form of life/living environment. Furthermore, the situated sensitivity to doing better or worse we refer to as 'situated normativity' is developed by taking part in practices.

Coordinating with others in a shared landscape of affordances

As we started out noting in section one, many of the skilled activities humans engage in are learned from others in their community. The affordances that invite an individual to act do so because of the individual's history of engaging in those practices, which gives them a sense for what it is appropriate to do in a particular situation. Inviting affordances thus contribute, albeit on a small scale, to the continuation of a wider practice. As an example of this dynamic consider how in the Netherlands, the trains used for long journeys will often have a few carriages in which travellers are expected to remain silent. The passengers sitting in a silent area can use their time on the train to work, enjoy reading a book, or listening to music without disturbance from other passengers. The continued availability of such affordances however depends on what train passengers do when they sit in such a place on the trains. Each passenger can make a small contribution to the continuation of this practice by refraining from answering their phone, or from entering into conversation with their fellow travellers. Their responding to this invitation contributes to the continuation of the affordances of this place. If each person were not to play their part, there would soon cease to be any difference between silent carriages and any other seat on the train (Van Dijk & Rietveld, 2017).

Whenever an individual responds appropriately to an inviting affordance in ways that are laid down in an ongoing practice, like in the example of the silent carriage, they are also coordinating with other members of this practice. The contribution to the continuation of the wider practice of silent areas is an example of action coordination albeit over longer time scales. An individual's participation in joint action can be compared with an individual's adequate participation in large-scale sociomaterial practices[10]. What these activities (acting jointly and acting in agreement with a standing practice) share in common is a sensitivity to doing better or worse in the particular situation in one's engagement with affordances. We call the multiple affordances that stand out as inviting the 'field of relevant affordances' (Bruineberg & Rietveld, 2014; Rietveld & Kiverstein, 2014). The affordances that belong to the field are those that contribute to improving an agent's grip on a situation. What counts as improvement in grip is however relative to one's sense of what actions the situation demands. In the silent train carriage for instance, receiving a phone call will invite declining if one is familiar with the customs of travelling by train in the Netherlands. For a tourist unfamiliar with these customs, a ringing phone will simply invite answering. The individual agent, in acting to improve grip, is acting based on their sense of what others in their social group or socio-cultural practice would do. We are suggesting that over time, this can be thought of in terms of acting in coordination with this socio-cultural practice or social group. What happens over this longer time scale isn't so different from what happens when individuals act together in the here and now.

So far, we have been concerned with the contribution of a field of relevant affordances to we-intentionality. In the next section we turn to the notion of an interpersonal synergy. We will see below that interpersonal synergies form through engagement with affordances of shared relevance. Joint actions within a dyad can also be thought of as a situated self-organising process of selectively responding to multiple relevant affordances at the same time. When the two individuals cooperate in the plank lifting experiment, they do so by adapting and adjusting their behaviour to each other and to the affordances of the plank. Importantly, Kerry Marsh (one of the experimenters involved in carrying out the original study) has compared this switch to what an individual does when she shifts from lifting an object with one hand to lifting a larger object with two hands (Marsh, 2015: 317–8). The process of switching from removing the plank as an individual agent to doing it together should be seen as a self-organising process that arises spontaneously, based on the dynamic coupling of the agents with the field of relevant affordances. It is in this respect no

[9] We borrow the term "sociomaterial practice" from anthropology and the field of organizational studies in which it is used to refer to the entanglement of social practice with the material forms of the places in which people act and interact (Mol, 2002; Orlikowski, 2007).

[10] In earlier work we have shown that the latter can be understood by the person's engagement with several relevant affordances simultaneously over multiple timescales. We'll see in the next section that the same readiness for multiple relevant affordances over different timescales can occur in cases of joint action.

different from the spontaneous switching when lifting an object with one hand, to lifting the object with two hands that happens within a subject based on the dynamics of the agent-environment coupling.

Interpersonal synergies

People coordinate and join forces in acting together by responding to affordances of shared relevance in the living environment. When an affordance is of shared relevance to two or more individuals, these agents will then be ready to act together, to cooperate with each other and to pool their individual abilities, so as to improve grip on the overall situation in which the affordances are nested. Moreover, the result of this pooling of their individual abilities is that actions become possible (like the lifting of the long plank) that were not possible for each of the participants acting as individuals[11].

Consider for instance a dyad using a two-handed saw to cut through a fallen tree. Person A takes control of the left side of the saw, while person B controls the right side of the saw. The affordances offered by the saw are to use a technical term, 'nested' in the wider context of the activity of sawing the piece of wood in order to build something together. Possibilities for action are often nested within other large-scale affordances (Van Dijk & Rietveld, 2021). The affordances of the saw in this example are embedded within the affordances of the shed where the building activity is taking place.

Dealing adequately with matters of shared relevance (like the liftability of the long plank) will require multiple individuals to combine their abilities. The result of this pooling of abilities is that affordances that are of shared relevance stand out from the landscape of affordances. Tending towards an improved grip on the situation is now something that individuals do in collaboration by acting together. They sense as a member of a dyad, or a larger group, how well or badly they are faring in their skilled engagement with the environment[12]. In the context of the planks experiment the participants can simply take it for granted that they are doing the experiment together with another cooperative participant prepared to do what the experimenter has asked of them. However, this is not something that can ordinarily be so readily assumed. If I am to embark on some joint endeavour with another agent, I must know that they are truly committed, and they must know the same about me. It might then be argued that for both of us to know this about each other will require each of us to first identify and ascribe intentions to the other, and perhaps even to know that we have done so.

How is it that the two agents managed to act in coordinated ways if they do not represent who is doing what, when and how? We suggest that the coordination of patterns of action readiness emerges across the individuals as a dyad in just the same way as it happens within each of them based on the self-organising dynamics of the agent-agent-environment system. We have seen a simple example of this self-organising dynamic in the plank lifting experiment. The two individuals act in coordinated ways because both are responding to a situation of shared relevance (on the basis of abilities that they have acquired thanks to a history of interactions in the same form of life). It is their mutual responsiveness to the same nested structure of affordances in the shared context of the experiment that accounts for how they are able to coordinate with each other. If a couple are visiting a DIY-store together as part of a shopping trip for a home renovation, they share a readiness for finding the items they need for the renovation, engaging with the shop-assistants to help them find said items, paying for them, and returning home[13]. Their activities are nested together in this way because the couple are each responsive to the large-scale affordance of the home renovation.

Similarly, in the plank lifting experiment, each agent is not only responsive to the affordance of liftability but to a whole nested structure situated in the living environment: to an entire field of relevant affordances. The readiness to lift the plank occurs in the wider context of the moving conveyor belt, as part of an experimental set-up in which the two subjects have been instructed. We are suggesting that each of these aspects (the plank, the moving conveyor belt, the experimental instructions) of the situation (and many others) should be thought of as providing relevant affordance that elicit a corresponding state of action readiness in each of the agents. The states of bodily action readiness are internal states of the individual that become coordinated with each other within the individual in such a way that they tend towards a grip on the nested structure of their field of relevant affordances as a whole. We can see this in the temporal dynamics of the individual's skilled behaviour. Each affordance-related state of action-readiness operates over different time scales. For instance, some states of action-readiness relate to the possible actions invited by the linguistic instructions of the experimenters, operating over longer time scales of the duration of the whole experiment. These slower states of action-readiness enslave the

[11] There is good reason to think that cooperating with others and pooling abilities played an important role in human evolutionary history, as has been convincingly argued by Tomasello, Sterelny and many others (Sterelny, 2012; Tomasello, 2014).

[12] When individuals form an interpersonal synergy do they, at the same time, form a group agent, or a first-person plural subject that is jointly committed to the performance of an activity together, for example going for a walk, or painting the house (see Gilbert, 1989; Gilbert, 2014; List & Pettit, 2011)? Something like this question comes when we consider whether a group agent forms out of the pooling of abilities that happens when interpersonal synergies form. If we agree that individuals that are acting as parts of an interpersonal synergy form a group agent, this would imply an important difference between skilled intentionality and skilled we-intentionality. In the former the skilled agent is an individual, while in skilled *we*-intentionality the agent is a dyad or a larger group. We think however that this difference is probably better understood in terms of time-scales. Over long time scales, individuals are always coordinating with the multiple members of practice. We saw this in the example of the silent train carriage. Over shorter time scales this coordination requires the pooling of abilities in ways that lead to the formation of interpersonal synergies. This is an idea we hope to return to on another occasion.

[13] See also our ethnography of a shopping trip to a carpet warehouse in van Dijk & Rietveld, 2021: p.351-2).

faster-changing states of action readiness that relate to the planks and to their movement on the conveyor belt.

Now it is not only the states of action readiness of the individuals that are coordinated with the nested structure of affordances in the field, and to the wider sociomaterial practices they share. Insofar as the two individuals are responding to relevant affordances, they will have similar or complementary bodily states of action readiness that are coordinated with each other. Instead of moving planks now imagine that a group of people are engaged in moving a heavy piano down some stairs. The movements of each person will have to complement each other if they are to avoid losing grip on the piano. Each person will however be moving in a slightly different way, some providing more of a supporting role for the piano while other will take more of a responsibility for moving the piano downwards. It is crucial that all the individual movements that are being made are carefully and tightly adjusted to each other. The affordances of the heavy piano play a crucial role in this process, as of course do those of the stairs the people are descending. The agents are continuously adjusting their movements to each other, but only in relation to the effect their respective movements are having on the moving of the piano in the wider context of the stairs. If the moving situation becomes so precarious that they feel they need an extra pair of hands, the possibility of ringing the neighbour's doorbell to ask them for help may become an affordance of shared relevance.

In joint action multiple individuals succeed in coordinating with each other because they form what we have called an interpersonal synergy. (We borrow this terminology from Dale *et al*., 2014). The movement scientist Nikolai Bernstein observed that the number of muscles and joints in the human body creates a problem of control (Bernstein, 1967; Kelso, 2009; Turvey, 1990). Each individual muscle and joint has multiple degrees of freedom. For each action the agent is ready to perform, there are very many different constellations of muscle and joint dynamics that can bring about the same action. Yet when we move our body, all of our muscles and joints work together in concert in ways that are somehow adjusted and fine-tuned to the affordances to which the agent is responding. If the individual is writing with a pencil for instance, this activity constrains the timing and force of contraction and relaxation of the individual muscles and joints in his hands, fingers, and arms. The joints and muscles appear not to act independently but interdependently, self-organising through their coupling, constraining each other's degrees of freedom, and thereby reducing the need for control. Similarly, when several people are moving a piano, it is the constraints that derive from people being coupled to the piano that reduce the degrees of freedom of their individual movements allowing them to function together as a single coordinated whole. The mutual adaptation makes the behaviour of the individuals interdependent in a way that makes it possible for them to pool their skills, so as to coordinate their responsiveness to multiple affordances of shared relevance. It is the affordances of the heavy piano in the situation of descending the steep stairs that invites the individuals carrying the piano together to coordinate their actions in a particular way. There is no need for them to plan who does what and when. Instead, this is settled on the fly through the interpersonal synergy they form as they couple to the piano in the particular local landscape of affordances encountered on the stairs.

In the final section, we pull together the two elements we have laid out so far of a field of relevant affordances containing affordances of shared relevance and interpersonal synergies to propose an account of we-intentionality in terms of skilled intentionality. We will see that interpersonal synergies form through engagement with affordances of shared relevance. These affordances constrain and limit the degrees of freedom in the behaviour of the interlocutors. The coordination that is achieved is best accounted for in terms of the behaviour of a single integrated system that is jointly shaped by the co-actors as they respond to the affordances of shared relevance on the basis of the skills they have developed by taking part in sociomaterial practices.

Skilled we-intentionality

In the previous section, we saw how it is the affordances of the joint liftability of the piano nested in the setting of the stairs around which the interpersonal synergy self-organises. The multiple nested affordances that make up the situation in which the piano is being moved together function as constraints on the self-organising process. Each relevant affordance generates micro-states of action readiness. Multiple simultaneous states of action readiness self-organise to form a macro-pattern of action readiness across the two individuals (an interpersonal synergy) in ways that are constrained by affordances of shared relevance. The affordances of the piano moving situation provide constraints on the behaviour of the individuals. Their skills attune them to this dynamically changing and evolving situation as they move the piano allowing them to maintain a good grip on the piano.

A central question for a theory of joint action, seen from this ecological dynamical perspective, is under what conditions affordances stand out as being of shared relevance inviting the coordinated actions of multiple individuals. We've seen (in section two with the example of the football player) how affordances stand out as inviting to an individual when they improve grip. One is ready to act on those relevant affordances that allow one to tend towards a better grip. The same is true when affordances are of shared relevance. Affordances stand out as of shared relevance when it is only by acting in ways that are mutually adapted that each of the individuals is able to improve grip. Suppose one is using a two-handed saw. This is the kind of saw that you can only use to cut through a fallen tree if when one agent pushes on the saw, the other agent pulls. The rhythm the two agents establish is thus crucial to using the saw skilfully. The two agents need to mutually adapt to each other producing complementary actions, just like in the case of the piano we just discussed. If they do not adapt to each other, they will have the sense as a dyad that what they are doing is not going well. So long as they mutually adapt each allowing their actions to be drawn from them by

the affordances of the two-handed saw, they will succeed in improving grip as a dyad. Thus, affordances will be experienced as shared invitations to multiple individuals when each of the individuals has the sense that a way to improve grip is by joining forces with other agents. We can suppose that this is a sense the individuals develop often through interaction with others. We learn that there are certain affordances such as those of two-handed saws that can only be engaged with skilfully by doing things together with other people.

We-intentionality, which is typically taken to be required for acting jointly, can therefore be explained in terms of affordances of shared relevance that act as constraints on the self-organising states of action readiness. The agents that participate in a joint action do not need to have predefined intentions or goals that specify who is to do what and when, in order to coordinate with each other. Multiple agents can have control over an action the states of action readiness self-organise to form interpersonal synergies in response to affordances of shared relevance.

It might be objected that people often do act based on explicit goals and intentions that they set for themselves in advance of acting[14]. The rich literature on planning to which Michael Bratman for instance has made important contributions is all about the cognitive dynamics of goal-setting (Bratman, 1987; cf. Castelfranchi & Paglieri, 2007; Pacherie, 2006). We have shown on the basis of ethnographic research over the course of many months in the practice of making art & architecture that even cases of long-term planning can naturally be described in terms of skilled intentionality (Rietveld *et al.*, 2022; van Dijk & Rietveld, 2017; van Dijk & Rietveld, 2020; van Dijk & Rietveld, 2021). The essence of what we found is that skilled agents, in this case visual artists and architects, are able to be responsive to large scale affordances, like the possibility to create an art installation that had never been made before and took five years to realize.

A simple example on a much shorter timescale might help to drive this point home. Suppose that I have agreed with my partner that I will cook tonight, and I am thinking about the ingredients I will need. The dish is one I've made many times before and because of my extensive practice I know exactly what to do to make it taste its best. I know in particular which vegetables and spices are needed. Yet, I don't have what I need and this situation makes the visit to the supermarket an attractive course of action. A supermarket visit is a possibility for action (a large-scale affordance) available in our human form of life and one with which I am very familiar. All of the thinking and planning in this example, can be readily made sense of in terms of my skilled engagement with this large-scale affordance over the time-scale of a few hours (Van Dijk & Rietveld, 2021). We would therefore resist making a distinction between emergent coordination and planned coordination as Knoblich *et al.* 2011 propose to do. We propose to understand all cases of coordination of action by multiple individuals as self-organized in terms of skilled intentionality. Examples of future-directed actions that seem to call for explanation in terms of who will do what and when, we propose to explain in terms of responsiveness to large-scale affordances such as visits to the supermarket that serve to organise the relevant affordances whose invitations weave together over time. Understanding action in terms of skilled intentionality allows for an explanation of how the individual is always ready for multiple possibilities that matter to them even when some of these possibilities are some way off in the future.

Consider, as a final example ostensive communication through pointing. Suppose we are out collecting fruit together and I point at some berries but all you can see are leaves. You can readily infer that I am not pointing at the leaves but most likely at some hidden berries. One might suppose that you can make this inference because the shared goal of finding berries is part of our common ground (Tomasello, 2014). There are many possible features of the current perceptual scene I might be attempting to direct your attention towards. To identify which of these features is actually of relevance to you, you must first determine what it is that I am intending to make you believe. You must think something along the following lines: his pointing in the direction of those leaves would make sense if he intends to make me believe there are berries hidden over there behind those leaves. Similarly, when I make the pointing gesture, I am also imagining how the gesture is likely to be received by you, and whether it is likely to be understood. I must think about what my partner will infer about my intentions, and the beliefs I intend to bring about in her. In contrast to this, we've argued that in perceiving the pointing gesture, you will be able to immediately recognise that I mean to draw your attention to the hidden berries, not to the leaves. It is the hidden berries that are of shared relevance to us in this situation. We are after all, engaged in the shared activity of hunting for berries. The pointing gesture is made in a situation

[14]Glenda Satne objected to us, in her review of an earlier version of this paper, that there will be many examples of we-intentionality that are not well described in terms of tending towards an optimal grip on a field of affordances of shared relevance. She gives as one example, helping an old lady to cross the road. She suggested that the system I temporarily form with the old lady is not one in which my grip on the road is improved. If anything it might be worsened by my act of helping her. We suggest however that the agent-agent-environment system one temporarily forms with the old lady is one that is more stable, and better adapted to the situation of crossing the road, than the system each of us forms when acting alone. Satne wonders how to make sense of other-oriented motivations like respect and kindness in our account? We suggested earlier that skills are always characterised in part by care and thoughtful sensitivity as part and parcel of the cycles of perception and action that couple the agent to the relevant affordances of the living environment. This care and thoughtfulness has its origin in a history of engaging in the practices of one's community. By virtue of this past history of acting, one has cultivated a sense that the right thing to do in this particular situation is to step in to help the lady across the road. It might be objected that it is better for the old lady but worse for me because my crossing becomes more difficult. However, this is to ignore the discontent I experience when I see the old lady at risk. Acting to help her allows me to improve grip because it reduces the feeling of discontent that is experienced as long as she appears to be at risk.

that is of shared relevance because we are undertaking this activity together.

Conclusion

Our aim in this paper has been to situate we-intentionality in relation to the social, cultural, material and natural environment, what we have called the 'living environment'. We have shown how people can succeed in acting together in coordinated ways without reasoning about each other's intentions. We have proposed an account of we-intentionality in terms of skilled intentionality, the coordination with multiple affordances simultaneously. When people coordinate with each other, they are responsive to (multiple) affordances of shared relevance. We have explained this responsiveness in terms of multiple states of bodily action readiness that self-organise in ways that are enabled and constrained by the affordances of shared relevance. Typically, people are able to coordinate with each other on the basis of skills for responding to affordances of shared relevance.

We've asked the concept of shared relevance to do a good deal of explanatory work in the account of joint action we have proposed. The key question in the account of joint action we propose is why it should be that an affordance (like the liftability of the plank) should stand out from the landscape of affordances as inviting or calling for the coordinated action of multiple individuals. Why should it be that individuals sometimes find themselves ready to join forces and collaborate in responding to affordances of shared relevance? The answer to this question is just the same as the answer we would give in the case of skilled intentionality for the actions of individuals. Affordances stand out as relevant for an individual agent because of disattunement in the agent-environment system as a whole, which the individual then acts to reduce. Affordances are of shared relevance, we have argued, when the only way to improve grip is by forming an interpersonal synergy with other individuals, like in the example of the two-handed saw. Affordances more generally stand out from the landscape as being of shared relevance to two or more individuals because of the sociomaterial practice of which the individuals are members. The berries stand out as relevant for the people that are foraging together for instance, because foraging is part of their shared way of life, and the berries are significant in the context of this shared activity.

Now it might be objected that all skilled intentionality turns out to be skilled-we intentionality. For whenever people act skilfully they are coordinating with other members of a practice. What this objection highlights is that there is indeed something common to the skilled actions of individuals, dyads and groups. What is common are the multiple relevant affordances and the shared sociomaterial practices with which individuals are coordinating when they act with skill. However, we have argued that there are also differences between what individuals do when they feel, think and act as dyads, or as members of larger groups. In the latter cases, the individuals pool their sensitivities to situated normativity, their abilities for distinguishing between better or worse ways of acting, so as to tend towards an improved grip on affordances of shared relevance as members of a dyad or group. The individuals are ready to join forces with relevant affordances and act together based on what matters to them as members of a practice.

Data availability

No data are associated with this article.

References

Abramova K, Slors M: **Social cognition in simple action coordination: a case for direct perception.** *Conscious Cogn.* 2015; **36**: 519–531.
PubMed Abstract | **Publisher Full Text**

Beer RD: **Dynamical approaches to cognitive science.** *Trends Cogn Sci.* 2000; **4**(3): 91–9.
PubMed Abstract | **Publisher Full Text**

Bernstein NA: **Coordination and Regulation of Movement.** New York: Pergamon Press, 1967.
Reference Source

Bratman ME: **Shared Agency: a Planning Theory of Acting Together.** Oxford: Oxford University Press, 2014.
Reference Source

Bratman ME: **Intentions, Plans and Practical Reason.** Cambridge, MA: Harvard University Press, 1987.
Reference Source

Bratman ME: **Shared intention.** *Ethics.* 1993; **104**(1): 97–113.
Reference Source

Bruineberg J, Rietveld E: **Self-organization, free energy minimization, and optimal grip on a field of affordances.** *Front Hum Neurosci.* 2014; **8**: 599.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Bruineberg J, Rietveld E: **What's inside your head once you've figured out what your head is inside of.** *Ecol Psychol.* 2019; **31**(3): 198–217.
Publisher Full Text

Bruner J: **Child's Talk.** New York, NY: Norton, 1983.

Butterfill SA: **Joint action and development.** *Philos Q.* 2012; **62**(246): 23–47.
Publisher Full Text

Castelfranchi C, Paglieri F: **The role of beliefs in goal dynamics: prolegomena to a constructive theory of intentions.** *Synthese.* 2007; **155**(2): 237–263.
Reference Source

Chemero A: **Sensorimotor empathy.** *J Conscious Stud.* 2015; **23**(5–6): 138–153.
Reference Source

Chemero A: **Radical Embodied Cognitive Science.** Cambridge, MA: MIT Press, 2009.
Reference Source

Dewey J: **Art as Experience.** New York: Minton Balch. 1934.
Reference Source

Dale R, Fusaroli R, Duran ND, *et al.*: **The self-organization of human interaction.** In B. Ross (Ed.) *The Psychology of Learning and Motivation.* Burlington MA,: Elsevier Academic Press, 2014; **54**: 43–95.
Publisher Full Text

Dreyfus HL: **Why Heideggerian AI failed and how fixing it would require making it more Heideggerian.** *Philos Psychol.* 2007; **20**(2): 247–68.
Publisher Full Text

Freeman WJ: **How Brains Make Up Their Minds.** New York: Columbia University Press, 2000.
Reference Source

Freeman WJ: **Simulation of chaotic EEG patterns with a dynamic model of the olfactory system.** *Biol Cybern.* 1987; **56**(2–3): 139–50.
PubMed Abstract | **Publisher Full Text**

Frijda NH: **The emotions.** New York: Cambridge University Press, 1986.
Reference Source

Frijda NH: **The Laws of Emotion.** Mahwah, NJ: Lawrence Erlbaum Associates, Inc. 2007.
Reference Source

Friston K: **The free energy principle: a unified brain theory?** *Nat Rev Neurosci.* 2010; **11**(2): 127–38.
PubMed Abstract | **Publisher Full Text**

Friston K: **Embodied Inference: or "I think therefore I am, if I am what I think".** *The implications of embodiment (Cognition and Communication).* 2011; 89–125.
Reference Source

Fusaroli R, Tylén K: **Carving language for social interaction: a dynamical approach.** *Interaction Studies.* 2012; **13**(1): 103–24.
Publisher Full Text

Gallagher S, Ransom T: **Artifacting minds: material engagement theory and joint action.** In C Tewes (Ed.) *Embodiment in Evolution and Culture.* Berlin: De Gruyter, 2016; 337–51.
Reference Source

Gibson JJ: **The Ecological Approach to Visual Perception.** New York: Psychology Press; 1979/2014.
Reference Source

Gilbert MP: **On Social Facts.** Princeton, NJ: Princeton University Press, 1989.
Publisher Full Text

Gilbert MP: **Living Together: Rationality, Sociality, and Obligation.** London: Rowman & Littlefield. 1996.
Reference Source

Gilbert M: **Joint Commitment: How We Make the Social World.** Oxford, UK: Oxford University Press. 2014.
Publisher Full Text

Haugeland J: **The intentionality all-stars.** *Philos Perspect.* 1990; **4**: 383–427.
Publisher Full Text

Heft H: **Ecological Psychology in Context: James Gibson, Roger Barker, and the Legacy of William James' Radical Empiricism**. New York: Taylor & Francis/ Psychology Press, 2001.
Publisher Full Text

Hobson P: **The Cradle of Thought.** Reading, UK: Palgrave Macmillan, 2002.
Reference Source

Käufer S, Chemero A: **Phenomenology – An Introduction**. Cambridge, UK: Polity Press, 2015.
Publisher Full Text

Kelso JAS: **Dynamic Patterns: The Self-Organisation of Brain and Behaviour**. Cambridge, MA: MIT Press, 1995.
Reference Source

Kelso JAS: **Synergies: atoms of brain and behavior.** *Adv Exp Med Biol.* 2009; **629**: 83–91.
PubMed Abstract | **Publisher Full Text**

Kiverstein J, Rietveld E: **Scaling-up skilled intentionality to linguistic thought.** *Synthese.* 2021; **198**: 175–194. Special Issue – Radical Theories of Cognition.
Publisher Full Text

Kiverstein J, van Dijk L, Rietveld E: **The field and landscape of affordances: Koffka's two environments revisited.** *Synthese: Special Issue Gestalt Psychology, Phenomenology and Embodied Cognitive Science.* 2019.
Publisher Full Text

Knoblich G, Butterfill S, Sebanz N: **Psychological research on joint action: theory and data**. In B Ross (Ed.) *The Psychology of Learning and Motivation.* Burlington MA,: Elsevier Academic Press, 2011; **54**: 59–101.
Publisher Full Text

Ladyman J, Wiesner K: **What is a Complex System?** London, UK: Yale University Press, 2020.
Publisher Full Text

List C, Pettit P: **Group Agency: The Possibility, Design, and Status of Corporate Agents.** Oxford, UK: Oxford University Press. 2011.
Publisher Full Text

Malafouris L: **How Things Shape the Mind: A Theory of Material Engagement**. Cambridge, MA: MIT Press, 2013.
Reference Source

Marsh KL: **Social ecological context of conversing.** *Ecol Psychol.* 2015; **27**(4): 310–334.
Publisher Full Text

Merleau-Ponty M: **The Structure of Behaviour**. Translated by Alden L Fisher, Boston: Beacon Press, 1943/1962.

Merleau-Ponty M: **Phenomenology of Perception**. Translated by Colin Smith, London: Routledge, 1945/2002.
Publisher Full Text

Mol A: **The Body Multiple: Ontology in Medical Practice**. Durham, NC: Duke University Press, 2002.
Publisher Full Text

Noë A: **Strange Tools: Art and Human Nature.** New York, NY: Hill & Wang. 2015.
Publisher Full Text

Orlikowski WJ: **Sociomaterial practices: Exploring technology at work.** *Organization Studies.* 2007; **28**(9): 1435–1448.
Publisher Full Text

Pacherie E: **Toward a dynamic theory of intentions**. In S Pockett, WP Banks, and S Gallagher (Ed's) *Does Consciousness Cause Behaviour? An Investigation of the Nature of Volition.* (Cambridge MA: MIT Press), 2006; 145–167.
Publisher Full Text

Pacherie E: **Is collective intentionality really primitive**. M Beaney, C Penco, & M Vignolo (Ed.'s) *Mental Processes: Representing and Inferring.* Cambridge, UK: Cambridge Scholar's Press, 2007; 153–175.
Reference Source

Rakoczy H: **The development of individual and shared intentionality.** In J Kiverstein (Ed.) *The Routledge Handbook of Philosophy of the Social Mind.* London: Routledge, 2017; 139–151.

Reddy V: **Why engagement?: A second-person take on social cognition**. In A Newen, L De Bruin & S Gallagher (Ed's) *The Oxford Handbook of 4E Cognition.* Oxford, UK: Oxford University Press, 2018; 433–453.
Publisher Full Text

Reed ES: **Encountering the world: Toward an ecological psychology**. New York, NY: Oxford University Press, 1996.
Reference Source

Reed ES, Bril B: **The primacy of action in development: A commentary on N. Bernstein**. *Dexterity and its Development.* (Translated, by ML Latash, Edited by ML Latash, MT Turvey), Mahwah NJ: Lawrence Erlbaum Associates Publishers, 1996; 431–451.
Reference Source

Reed ES: **Cognition as the cooperative appropriation of affordances.** *Ecol Psychol.* 1991; **3**(2): 135–158.
Publisher Full Text

Richardson MJ, Campbell WL, Schmidt RC: **Movement interference during action observation as emergent coordination.** *Neurosci Lett.* 2009; **449**(2): 117–122.
PubMed Abstract | **Publisher Full Text**

Richardson MJ, Marsh KL, Schmidt RC: **Challenging the egocentric view of coordinated perceiving, acting and knowing**. In LF Barrett, B Mesquita, & ER Smith (Ed's) *Mind in Context.* New York: Guilford Press, 2010; 307–33.
Reference Source

Richardson MJ, Marsh KL, Baron RM: **Judging and actualizing intrapersonal and interpersonal affordances.** *J Exp Psychol Hum Percept Perform.* 2007; **33**(4): 845–859.
PubMed Abstract | **Publisher Full Text**

Rietveld E: **Situated Normativity: The Normative Aspect of Embodied Cognition in Unreflective Action.** *Mind.* 2008; **117**(468): 973–1001.
Publisher Full Text

Rietveld E, Denys D, van Westen M: **Ecological-enactive cognition as engaging with a field of relevant affordances: The skilled intentionality framework (SIF)**. In A Newen, L De Bruin, & S Gallagher (Eds.), *The Oxford handbook of 4E (embodied, embedded, extended, enactive) cognition.* Oxford: Oxford University Press, 2018; 41–70.
Publisher Full Text

Rietveld E, Kiverstein J: **A rich landscape of affordances.** *Ecol Psychol.* 2014; **26**(4): 325–352.
Publisher Full Text

Rietveld E, Rietveld R, Martens J: **Collaborative embodied engagements with strategic interventions: a conversation.** In K. Bicknell & J. Sutton (Ed's) *Collaborative Embodied Performance: Ecologies of Skill.* London, UK: Methuen Drama, Bloomsbury Publishing. 2022.

Rochat P, Robbins E: **Sharing and fairness in development.** In J. Kiverstein (Ed.) *The Routledge Handbook of Philosophy of the Social Mind.* London: Routledge, 2017; 222–243.
Publisher Full Text

Rogoff B: **Apprenticeship in Thinking: Cognitive Development in Social Context**. Oxford, UK: Oxford University Press, 1990.
Reference Source

Satne G: **Interaction and self-correction.** *Front Psychol.* 2014; **5**: 798.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Satne G, Sallice A: **Shared intentionality and the cooperative evolutionary hypothesis.** En. A. Fiebich (Ed.). *Minimal Cooperation and Shared Agency.* New York, NY: Springer, 2020; 71–92.
Publisher Full Text

Searle JR: **Collective intentions and actions**. In P. Cohen, J. Morgan & M.E. Pollack (Ed's) *Intentions in Communication.* Cambridge: Cambridge University Press, 1990; 90–105.
Reference Source

Searle JR: **The Social Construction of Reality**. New York: Free Press. 1995.
Reference Source

Segundo-Ortin M, Satne G: **Sharing attention, sharing affordances: from dyadic interaction to collective information.** In D. D'Angelo, M. Merle & E. Solomonova (Ed's) *Mediation and Access. A New Approach to Attention*. De Gruyter: Berlin. (forthcoming).
Reference Source

Sellars W: **Science, Perception and Reality**. London: Routledge & Kegan Paul. 1963.
Reference Source

Shockley K, Richardson DC, Dale R: **Conversation and coordinative structure.** *Top Cogn Sci.* 2009; **1**(2): 305–319.
PubMed Abstract | **Publisher Full Text**

Sterelny K: **The Evolved Apprentice.** Cambridge, MA: MIT Press. 2012.
Reference Source

Strogatz SH: **Non-Linear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering**. New York, NY: Perseus Books. 1994.
Reference Source

Suchman LA: **Plans and Situated Actions: the Problem of Human Machine Interaction**. Cambridge: Cambridge University Press. 1987.
Reference Source

Tollefsen D, Dale R: **Naturalising joint action: a process-based approach.** *Philos Psychol.* 2012; **25**(3): 385–407.
Publisher Full Text

Tomasello M: **A Natural History of Human Thinking**. Cambridge, MA: Harvard University Press. 2014.
Reference Source

Tschacher W, Haken H: **Intentionality in non-equilibrium systems? The functional aspects of self-organised pattern formation.** *New Ideas Psychol.* 2007; **25**(1): 1–15.
Publisher Full Text

Tsuda I: **Towards an interpretation of dynamic neural activity in terms of chaotic dynamical systems.** *Behav Brain Sci.* 2001; **24**(5): 793–809.
PubMed Abstract | **Publisher Full Text**

Tuomela R: **We-intentions revisited.** *Philosophical Studies.* 2005; **125**(3): 327–369.
Publisher Full Text

Tuomela R, Miller K: **We-intentions.** *Philosophical Studies.* 1988; **53**: 367–389.
Publisher Full Text

Turvey MT: **Coordination.** *The American Psychologist.* 1990; **45**: 938–953.
Publisher Full Text

Van Den Herik J, Rietveld E: **Reflective situated normativity.** *Philosophical Studies.* 2021.
Publisher Full Text

Van Dijk L, Rietveld E: **Situated imagination.** *Phenomenology & Cognitive Sciences.* 2020.
Publisher Full Text

Van Dijk L, Rietveld E: **Situated anticipation.** *Synthese.* 2021; **198**(1): 349–371.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Van Dijk L, Rietveld E: **Foregrounding sociomaterial practice in our understanding of affordances: the skilled intentionality framework.** *Front Psychol.* 2017; **7**: 1969.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Vincini S, Gallagher S: **Developmental phenomenology: examples from social cognition.** *Cont Philos Rev.* 2021; **54**(2): 183–199.
Publisher Full Text

Vygotsky LS: **Mind in Society: The Development of Higher Psychological Processes.** (Ed.'s M Cole, VJ Steiner, S Scribner, & E Souberman), Cambridge, MA: Harvard University Press. 1978.
Publisher Full Text

Withagen R, de Poel HJ, Araujo D, *et al.*: **Affordances can invite behavior: Reconsidering the relationship between affordances and agency.** *New Ideas Psychol.* 2012; **30**(2): 250–258.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 2

Reviewer Report 15 October 2021

https://doi.org/10.21956/openreseurope.15233.r27687

© **2021 Satne G.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Glenda Satne**
Philosophy, School of Humanities and Social Inquiry, Faculty of Arts, Social Sciences and Humanities, University of Wollongong, Wollongong, Australia

I have reviewed the authors replies and the updated paper, I am happy to accept the paper as it is, without further revisions.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 05 October 2021

https://doi.org/10.21956/openreseurope.15233.r27686

© **2021 Chemero A.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Anthony Chemero**
Departments of Philosophy and Psychology, University of Cincinnati, Cincinnati, OH, USA

The authors have addressed my concerns. I approve this paper for indexing.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 22 July 2021

https://doi.org/10.21956/openreseurope.14481.r26951

© **2021 Chemero A.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Anthony Chemero
[1] Departments of Philosophy and Psychology, University of Cincinnati, Cincinnati, OH, USA
[2] Departments of Philosophy and Psychology, University of Cincinnati, Cincinnati, OH, USA

Comments on Kiverstein and Rietveld, “Skilled We-intentionality”, 2021

This paper is a valuable notification for philosophers who work on social and group cognition that ecological and dynamical psychologists have been building up a substantial body of work that is directly relevant to their concerns. These psychologists have built up an expanding body of empirical results and theoretical arguments that explain joint action without any reference to mental representations—whether individual, joint, or aligned. Kiverstein and Rietveld argue that shared affordances structure joint actions, without representations. This is in direct contradiction to the assumptions that have structured the philosophical debate. Because Kiverstein and Rietveld cite all the key players in the philosophical debate, those key players are much more likely to pay some attention than they have. Most of us in the ecological and dynamical communities typically ignore these philosophers; and they ignore us right back. (I have been guilty of this in my own writing on joint action.) This paper is a worthwhile contribution and worth publishing for that reason alone.

That said, I do have one major concern about the role of the Skilled Intentionality Framework in the argument. Let me preface the concern by saying that I am a fan of the Skilled Intentionality Framework and am even an author on one of the papers promoting and elaborating it. However, some of my colleagues in the ecological and dynamical psychology community complain that the Skilled Intentionality Framework is primarily a re-branding of ideas that were already in the scientific literature, and that its only innovations are terminological (and so unnecessary). I do not share this complaint, and defend the Skilled Intentionality Framework when it is raised. But it does align with my main concern with this paper. How much argumentative work is the Skilled Intentionality Framework actually doing in this paper?

The main argumentative structure of the paper goes from affordances, via development, to shared affordances to social coordination dynamics to interpersonal synergies to self-organized joint action on shared affordances, which obviates the need for mental representations in the explanation of joint action. This is a compelling argument. (In fact, I am finishing a book that follows some of the same path.) But notice that I did not mention the Skilled Intentionality Framework. What does the Skilled Intentionality Framework add to the mix? More importantly, what does the Skilled Intentionality Framework add that isn’t already there from ecological psychologists like Reed and Heft and Turvey and Shaw?

In my opinion, the main contribution of the Skilled Intentionality Framework here is the distinction affordances and invitations, but that distinction doesn't get much discussion. Affordances are opportunities for action; invitations are those opportunities for action that align with current goals and motivations. Kiverstein and Rietveld draw on Merleau-Ponty's idea of motor intentionality, which they call 'skilled intentionality' to bridge this gap between what we can do and what we might actually do. A second smaller contribution of the Skilled Intentionality Framework is the story it tells about the shared maintenance of cultural affordances. The latter of these is already available in the work of Reed and Heft that they cite. The former might be available in Shaw's work on intentional dynamics which they don't cite (and which I am not sure I understand). This leads to the worry that the primary contribution of the Skilled Intentionality Framework really is terminological in this case. What Kiverstein and Rietveld have done, according to this worry, is used some new terminology ('skilled intentionality', 'the landscape of affordances', 'the field of affordances') to set out some old ideas.

As I said above, this paper is worthwhile as is despite this concern just for the likelihood that it will lead to some cross-pollination between philosophers of joint action and ecological psychologists. But I do think that Kiverstein and Rietveld should make much more clear exactly what the Skilled Intentionality Framework is adding to the argument, and differentiating between that and what is already available in the ecological and dynamical psychology literatures. This is important not just for this paper, but for the Skilled Intentionality Framework more generally.

Is the work original in terms of material and argument?
Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?
Yes

Is the work clearly and cogently presented?
Yes

Is the argument persuasive and supported by evidence?
Yes

If any, are all the source data and materials underlying the results available?
No source data required

Does the research article contribute to the cultural, historical, social understanding of the field?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: philosophy of cognitive science, ecological psychology, dynamic systems theory

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 08 Sep 2021

Julian Kiverstein

Clarifying our aims

The aim of this paper was to connect research on social coordination in ecological psychology with our theoretical research on affordances, along the same lines as an earlier publication (Van Dijk & Rietveld 2017). We have argued in previous work for a theory that defines the affordances of the human environment in relation to a shared social, cultural and material form of life. This work has been inspired by among others Harry Heft and Ed Reed (Heft 2001; Reed 1996). Some of the research on social coordination that has been done in ecological psychology (e.g. Richardson *et al*. 2007) has been taken up in the philosophical literature on joint action by philosophers like Stephen Butterfill (e.g. Knoblich *et al* 2011) and Deborah Tollefsen (e.g. Tollefsen & Dale 2011). However, these philosophers have also tended to insist on a distinction between a form of social coordination which is emergent, and well-described by ecological psychology, and a form of joint action that is planned and calls for a more traditional cognitive explanation in terms of shared intentions and goals. More specifically, we tried to bring our research on affordances to bear to show that such a distinction between planned and emergent coordination is not needed. This led us to think about how to generalise our work on skilled intentionality, which has up until now focused on how the skilled individual is ready to respond to multiple relevant affordances at the same time, to thinking about skilled actions performed by dyads or groups. We wanted to show how the concept of skilled intentionality can be applied to a domain of action typically thought to call for higher-cognition, namely we-intentionality. Indeed, we-intentionality is sometimes argued to be associated with distinctively human forms of thinking and acting – see, for example the work of Michael Tomasello (2019). Our aim was to show how SIF might be able to make sense of what is according to some a uniquely human form of thinking and acting. If such a project succeeds it would correct for an often heard complaint that research in ecological-enactive cognitive science while well-suited for understanding perception and action will not scale up to explain what is distinctive about human cognition (see also Rietveld & Kiverstein 2014; Kiverstein & Rietveld 2018, 2020).

Competing Interests: No competing interests were disclosed.

Author Response 08 Sep 2021

Julian Kiverstein

2. **Is SIF largely "an exercise in rebranding"?**

One of the core concepts of our paper – interpersonal synergies – is indeed borrowed from the literature in ecological psychology on dynamics of social interaction in conversation (Dale *et al*. 2014). The novel contribution of our paper is however to describe how interpersonal synergies self-organise in relation to the field of relevant affordances. This is

not an idea that is present in the earlier literature but it is one that opens up to us because of the distinction we make between affordances and invitations based on the phenomenology of skilled action, and between the landscape of available affordances and the field of relevant affordances based on our own earlier work on the philosophy of affordances. Of course, both the concept of affordances and that of invitations were also present in the phenomenological literature prior to our developing SIF. The distinction between the landscape of affordances and the field of relevant affordances was introduced by us. As we explained at the beginning of our reply, we developed SIF to provide bridging concepts that could connect the phenomenological literature to work in ecological dynamics and neurodynamics.

Competing Interests: No competing interests were disclosed.

Author Response 08 Sep 2021

Julian Kiverstein

Large-scale affordances and temporalising affordances:
Finally, our recent work on large-scale affordances, and the temporalizing of the landscape of affordances, is an innovation of SIF made possible by our definition of affordances in relation to sociomaterial practices. The idea of skilled agents engaging with large-scale affordances is important in this article for resisting the distinction that Butterfill and colleagues have made between 'emergent" and "planned" coordination. In her review of this paper, Glenda Satne suggested we should restrict the ambitions of our paper to emergent coordination. She objected that to attempt to generalise the concept of skilled intentionality to planned coordination was stretching this concept beyond its usefulness. However, we have made explicit in our reply to her above why we disagree and prefer to remain more ambitious. The landscape of affordances in the human form of life includes possibilities for actions that traditionally would have been characterized as forms of 'higher' cognition. On the basis of ethnographic work we have shown that this richness of the landscape allows for an account to be given of complex team activities such as creating jointly a building or artwork that does not yet exist and situated talking in terms of skilled intentionality (van Dijk & Rietveld, 2018, 2020; van den Herik & Rietveld 2020).Understanding skilled action in terms of responsiveness to multiple relevant affordances over time, including large-scale affordances, allows us to account for both emergent and planned coordination. These are important theoretical contributions of SIF that should help with the scaling-up objection often levelled against work in ecological dynamics and embodied cognition.

Competing Interests: No competing interests were disclosed.

Author Response 08 Sep 2021

Julian Kiverstein

The main aim of the Skilled Intentionality Framework (SIF) is to provide an integrative conceptual framework allowing for connections to be made between otherwise disparate research fields of philosophical phenomenology, ecological dynamics, enaction,

neurodynamics, and anthropology. The aim of this paper was to connect research on social coordination in ecological psychology with our theoretical research on affordances, along the same lines as an earlier publication (Van Dijk & Rietveld 2017). We have argued in previous work for a theory that defines the affordances of the human environment in relation to a shared social, cultural and material form of life. This work has been inspired by among others Harry Heft and Ed Reed (Heft 2001; Reed 1996). Some of the research on social coordination that has been done in ecological psychology (e.g. Richardson et al. 2007) has been taken up in the philosophical literature on joint action by philosophers like Stephen Butterfill (e.g. Knoblich et al 2011) and Deborah Tollefsen (e.g. Tollefsen & Dale 2011). However, these philosophers have also tended to insist on a distinction between a form of social coordination which is emergent, and well-described by ecological psychology, and a form of joint action that is planned and calls for a more traditional cognitive explanation in terms of shared intentions and goals. More specifically, we tried to bring our research on affordances to bear to show that such a distinction between planned and emergent coordination is not needed. This led us to think about how to generalise our work on skilled intentionality, which has up until now focused on how the skilled individual is ready to respond to multiple relevant affordances at the same time, to thinking about skilled actions performed by dyads or groups. We wanted to show how the concept of skilled intentionality can be applied to a domain of action typically thought to call for higher-cognition, namely we-intentionality. Indeed, we-intentionality is sometimes argued to be associated with distinctively human forms of thinking and acting – see, for example the work of Michael Tomasello (2019). Our aim was to show how SIF might be able to make sense of what is according to some a uniquely human form of thinking and acting. If such a project succeeds it would correct for an often heard complaint that research in ecological-enactive cognitive science while well-suited for understanding perception and action will not scale up to explain what is distinctive about human cognition (see also Rietveld & Kiverstein 2014; Kiverstein & Rietveld 2018, 2020).

So, is this an exercise in “rebranding” ideas that are already in the scientific literature in the terms of SIF? One of the core concepts of our paper – interpersonal synergies – is indeed borrowed from the literature in ecological psychology on dynamics of social interaction in conversation (Dale et al. 2014). The novel contribution of our paper is however to describe how interpersonal synergies self-organise in relation to the field of relevant affordances. This is not an idea that is present in the earlier literature, but it is one that opens up to us because of the distinction we make between affordances and invitations based on the phenomenology of skilled action, and between the landscape of available affordances and the field of relevant affordances based on our own earlier work on the philosophy of affordances. Of course, both the concept of affordances and that of invitations were also present in the phenomenological literature prior to our developing SIF. The distinction between the landscape of affordances and the field of relevant affordances was introduced by us. As we explained at the beginning of our reply, we developed SIF to provide bridging concepts that could connect the phenomenological literature to work in ecological dynamics and neurodynamics. Finally, our recent work on large-scale affordances, and the temporalizing of the landscape of affordances, is an innovation of SIF made possible by our definition of affordances in relation to sociomaterial practices. The idea of skilled agents engaging with large-scale affordances is important in this article for resisting the distinction that Butterfill and colleagues have made between ‘emergent” and “planned” coordination.

In her review of this paper, Glenda Satne suggested we should restrict the ambitions of our paper to emergent coordination. She objected that to attempt to generalise the concept of skilled intentionality to planned coordination was stretching this concept beyond its usefulness. However, we have made explicit in our reply to her above why we disagree and prefer to remain more ambitious. The landscape of affordances in the human form of life includes possibilities for actions that traditionally would have been characterized as forms of 'higher' cognition. On the basis of ethnographic work, we have shown that this richness of the landscape allows for an account to be given of complex team activities such as creating jointly a building or artwork that does not yet exist and situated talking in terms of skilled intentionality (van Dijk & Rietveld, 2018, 2020; van den Herik & Rietveld 2020). Understanding skilled action in terms of responsiveness to multiple relevant affordances over time, including large-scale affordances, allows us to account for both emergent and planned coordination. These are important theoretical contributions of SIF that should help with the scaling-up objection often levelled against work in ecological dynamics and embodied cognition.

References Bratman, M.E. 1993. Shared intention. Ethics 104: 97-113. Bruineberg, J. & Rietveld, E. 2019. What's inside your head once you've figured out what your head is inside of. Ecological Psychology 31 (3): 198-217. Frijda, N.H., 2007. The Laws of Emotion, Mahwah, NJ: Lawrence Erlbaum Associates, Inc. Gallagher, S. & Zahavi, D. 2008. The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science. (London and New York, NY: Routledge, 2008). Heft, H. 2001. Ecological Psychology in Context: James Gibson, Roger Barker, and the Legacy of William James' Radical Empiricism. New York: Taylor & Francis/Psychology Press.. Kaufer, S. & Chemero, A. 2015. Phenomenology: An Introduction. Cambridge, UK: Polity Press. Knoblich, G., Butterfill, S., & Sebanz, N. 2011. Psychological research on joint action: theory and data. In B. Ross (Ed.) The Psychology of Learning and Motivation Vol. 54. Burlington MA,: Elsevier Academic Press, pp.59-101. Reed, E. 1996. Encountering the world: Toward an ecological psychology. New York, NY: Oxford University Press. Richardson, M.J., Marsh, K.L., & Baron, R.M. 2007. Judging and actualising intrapersonal and interpersonal affordances. Journal of Experimental Psychology: Human Perception and Performance. 33: 845-859. Rietveld, E., & Kiverstein, J. (2014). A rich landscape of affordances. Ecological Psychology, 26(4), 325–352. doi:10.1080/10407413.2014.958035. Rietveld, E. (2008). Situated Normativity: The Normative Aspect of Embodied Cognition in Unreflective Action. Mind. Tollefsen, D. & Dale, R. 2011. Naturalising joint action: a process-based approach. Philosophical Psychology 25 (3): 385-407. Tomasello, M. 2019. Becoming Human: A Theory of Ontogeny. Cambridge, MA: Harvard University Press. Van Den Herik, J. & Rietveld, E. 2020. Reflective situated normativity. Philosophical Studies Van Dijk, L. & Rietveld, E. 2017. Foregrounding sociomaterial practice in our understanding of affordances: the skilled intentionality framework.

Competing Interests: No competing interests were disclosed.

Reviewer Report 08 July 2021

https://doi.org/10.21956/openreseurope.14481.r26954

© 2021 Satne G. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Glenda Satne
[1] Philosophy, School of Humanities and Social Inquiry, Faculty of Arts, Social Sciences and Humanities, University of Wollongong, Wollongong, Australia
[2] Philosophy, School of Humanities and Social Inquiry, Faculty of Arts, Social Sciences and Humanities, University of Wollongong, Wollongong, Australia

"Skilled we-intentionality: Situating joint action in the living Environment" argues that the literature on joint action, mainly devoted to understanding different activities in which two or more individuals act together, rather than alone or in parallel, has so far underplayed the role of environmental affordances in accounting for joint activities. To address this omission, the paper proposes a development of the already existing framework of skilled intentionality to study cases of we-intentionality in which individuals are engaged with others in joint activities.

The paper is a good contribution to the literature and develops an interesting account of skilled action that incorporates the key role of the living environment in understanding joint action. Yet, the main claims in the paper are not entirely clear, and some of the key concepts proposed in the paper are ambiguous (in particular, that of 'grip' and 'shared relevance'). This is because the paper is covering a variety of cases of joint activity that seem to work differently. Consequently, the key concepts put forward in the paper appear to apply differently to different cases of joint activity, if they apply at all. The paper needs to state its ambitions and main aim more clearly, as well as refine some of its central claims and conclusions. In what follows, I substantiate these points in more detail, while also commenting on the claims in the paper that invite clarification or expansion, for doing so can improve an already interesting and worthwhile paper.

1.The first observation concerns the key concepts advanced in the paper.

In the following passage we are given a general characterization of the view:

"People coordinate and join forces in acting together by responding to affordances of **shared relevance** in the living environment. When an affordance is **of shared relevance** to two or more individuals, these agents will then be ready to act together, to cooperate with each other and to pool their individual abilities, so as to **improve grip** on the overall situation in which the affordances are nested." (p.7, my emphasis)

According to this characterization, joint action coordination takes place when the environment offers the opportunity/invitation for individuals to act together. This is, when affordances are of 'shared relevance' for the individuals (ready to be) involved in the action. 'Shared relevance' is defined in terms of 'improvement of grip", and 'improvement of grip' for a group of agents is defined in analogy with the individual case, that is, in terms of the reduction of disattunement in the agent-environment system**:**

"Why should it be that individuals sometimes find themselves ready to join forces and collaborate in responding to affordances of shared relevance? **The answer to this question is just the same**

as the answer we would give in the case of skilled intentionality for the actions of individuals. Affordances stand out as **relevant** for an individual agent because of **disattunement in the agent-environment system, which the individual then acts to reduce.** [...].” (p.10, my emphasis)

Coordination is then characterized as a form of “interpersonal synergy” (pp. 7-8), in terms of which “improvement of grip” is jointly obtained. Thus, circumstances in which affordances stand out as of ‘shared relevance’ are those in which “the **only way** to improve grip is by forming an interpersonal synergy with other individuals, like in the example of the two-handed saw.” (p.10, my emphasis).

It is important to note that the authors define a situation where there is shared relevance of affordances when “the only” way of improving grip is by forming an interpersonal synergy. Yet this description applies to certain forms of shared activity and not others.

Some of the examples the authors give such as moving a heavy piano together down the stairs (p.8), or joining another agent to use a two-handed saw to cut through a tree (p.7), are cases where the definitions just discussed seem to apply cogently. Yet, in many cases of shared activity, there is no clear ‘grip’ improvement that is taking place. For example, imagine a crowd in a stadium joining in a ‘Mexican wave’ to celebrate the victory of their soccer team and someone sitting in the same stand as these winning team supporters, but who is not a supporter of that team and does not join in the wave. Later, her own team scores a goal and she celebrates by singing a song that the supporters of her team are singing at the other side of the stadium. How should we think of her not joining the wave in terms of the view proposed? Should we say that she is not reducing the disattunment in the agent-environment system by not joining in or should we say the opposite? What about the singing in the other side of the field? There is no obvious explanation in terms of disattunement here, for there seem to be different dimensions of explanation involved. The person’s relation to her own team, and her determination to act as a member of that team, might say more about her motivations and actions than her bodily situation in the current context.

Other sets of cases that seem very different from the two-handed saw are cases that involve acting with others, but that are not cases of bodily interaction. For example, we all paint the house together, but we came to paint on different days with explicit instructions distributed to us in advance. In this case, it seems that we need other elements apart from interpersonal synergy to explain coordination between agents. Though, in this case, the environment plays a key role, there is much more to be said about how these relations work - between agents, between agents and environment, in relation to plans, etc. - compared to the double hand-saw case.

Another interesting kind of case is the following: I am about to cross the street and I see an old lady struggling to step down the sidewalk. I step in to help her. Maybe I do not clearly decide in the sense of making a conscious decision to help her, rather I jump in and help her cross the street. This action of mine arguably reduces my attunement with the environment for I am now part of a dyad with the old lady, and we, crossing together the street, are much clumsier and slower as a joint agential system that I would be if crossing alone (though, arguably, she is doing better in this joint action scenario than alone). It is not clear what the affordances of shared relevance are in this case, for I do not improve my grip on the situation by joining her, though when we are in fact crossing the street some degree of reduction of disattunment might help

explain - in part - our bodily coordination. It may be argued that, in this case, it is my values and the prevailing social norms by which I orient my action that explain what I am doing and why I am doing it, and perhaps even how do it (since I am gentle, I wait longer than I would with a younger person, I do this because of respect and kindness, even if perhaps this is not the most efficient way for us to cross the street (could I carry her in my arms instead?)). The authors come close to this sort of consideration when they say:

"Affordances more generally stand out from the landscape as being of shared relevance to two or more individuals because of the socio-material practice of which the individuals are members (There is something that the individuals care about because of the practices in which they take part, and they are ready to join forces with relevant affordances and act together based on what matters to them as members of a practice.)" (p.9)

This might be true. Incidentally this explanation seems to fit better not only with my three own cases here, but also with their own examples of the 'silent carriage in the train' (p.7) and those of 'shopping for a home-renovation' (p.8) and 'hunting for berries together' (p.9), in which a shared activity is explain in terms of other shared activities and practices. But it is to be noted how these examples, as much as mine above, are hardly explainable in terms of the notions of improvement of grip and reducing disattunement in the agent-environment system, as defined in the paper.

The disanalogy here runs deep, for if we were to keep the same terminology to explain how individuals are motivated to act and coordinate in all these cases, the terms would become highly metaphorical. Can my sitting silent in a silent carriage of a train be completely explained in terms of improvement of grip, independently of my aim of respecting the social norm of 'keeping quiet' that prevails in that setting? Can we explain Mexican waves or helping an old lady cross the street as agents' ways of improving their grip in the agent-environment system or do we need in addition, and crucially, reference to the agents' intentions, values or group memberships? These are cases in which socio-cultural norms and values in place partly explain what the agents are doing and why. In this, these cases are quite different than handling a piano together or a double-handed saw. In the latter, there is a clear bodily sense in which the action involves two agents, and its success requires bodily coordination that does indeed improve grip in the agents-environment system – the interpersonal synergy the authors speak about. Moreover, these same cases seem explainable in terms of coordination, independently of any further aims that the agents may have, except those that the coordination brings about, i.e. cutting the tree or moving the piano down the stairs. In the examples I provided though, whereas in some cases there is bodily coordination (e.g., the stadium wave; helping an old lady cross the street), this is either not a straightforward case of improving grip, and/or the motivations of the agents are not completely explainable independently of further values, intentions and aims. For these kinds of cases, we cannot explain what the affordances of shared relevance are merely in terms of improvement of grip, where this is understood in terms of reduction of disattunment of the agent-environment system by the sole means of forming interpersonal synergies. To explain all these cases in that way would turn terms like "relevance", "disattunement", and "grip" highly metaphorical and vague, for they would explain bodily coordination and sparse time-space coordination, cases where there is face-to-face interaction and cases where agents are anonymous to each other, cases of joint action motivated by group membership, and cases motivated by emotional contagion, etc., in exactly in the same way. It looks as if we do need more nuances here.

2. Another key issue that results from the different cases that the authors aim to explain by

the same means, is that depending on what case we focus on, the dialectics with the current accounts of joint action changes.

In p.9, the authors claim:

"The agents that participate in a joint action **do not need to have predefined intentions or goals that specify who is to do what and when, in order to coordinate with each other**. Multiple agents **can have control over an action the states of action readiness self-organise to form interpersonal synergies in response to affordances of shared relevance**."

While this is true in the case of the two-handed saw, it is not true of the case of helping the old lady, or in their own examples of "hunting for berries" or "shopping for a home renovation". In these cases, it seems that we need further explanation of what the agents are doing and those involve reference to something more: some goals, intentions, values, or social norms that the agents share. In this respect, these cases are sufficiently different from cases which can be understood in terms of "interpersonal synergies in response of affordances of shared relevance".

(NB: Some authors have attempted systematic characterizations and classifications of the different cases of joint action (see e.g. Knoblich et al. 2011 and Satne 2020a). According to these classifications, some joint/shared activities involve shared planes and/or shared intentions while others can be explained without making reference to any prior intentions, goals or plans (see Satne 2020b and Butterfill 2012) for alternative accounts of how to characterize each case). The latter are described as forms of "emergent coordination" (Knoblich et al. 2011), "minimalist" forms of joint action, (Butterfill 2016) or "basic joint intentional activities" (Satne 2020a, b; Fiebich & Gallagher 2013).)

This important difference between different kinds of joint activity has a direct impact on the claim that participants in joint action "do not need to have predefined intentions **that specify who is to do what and when, in order to coordinate with each other".** (p.9). When we focus on cases of joint action that are not defined by previous intentions or goals, in which bodily coordination plays a key role, i.e. the two-handed saw, or moving a piano together up or down the stairs, the authors' argument runs smoothly, and the claim that their account can do without attributing intentions to explain coordination is sound (for a similar claim see Satne 2020a and Knoblich 2011). Thus, for these cases the concept of 'interpersonal synergy' and that of affordances of 'shared relevance' do indeed explain the kind of bodily coordination between different agents that takes place and how agents might be motivated to participate in joint action without having any predetermined intentions or goals. But when we consider the larger set of cases that they use to illustrate the view, this simply does not follow. 'Shopping for a home-renovation', 'being in the silent carriage', 'painting a house together at different times', they all involve some pre-determined goals, intentions or norms in terms of which who takes part in a joint action and when are determined. Even if the authors would want to resist this claim, much more argument would be needed. Crucially this depends on how one understands what intentions, goals, or norms are, and how they work. Thus, much more argument would be required to substantiate the claim that "interpersonal synergies in response of affordances of shared relevance" could explain such cases - and every possible case of shared activity - in the absence of any intentions, plans, goals, or norms.

In sum, there are key issues that the authors need to address to clarify their proposal and make it cogent. Depending on how they address these, they could follow two alternative paths:

1. circumscribe the focus of the paper to explain bodily coordination in joint action. This can include cases of 'emergent coordination' (Knoblich et al. 2011), that is, cases of bodily coordination where there are no previous plans or preceding intentions or goals. This study can also include as one kind of case of bodily coordination, instances of acting together that involve bodily coordination but also further aims or goals (e.g. helping the old lady, riding on the silent carriage), in as far as it is clearly stated that the paper's aim is not to explain all aspects of these joint actions, but only bodily coordination in relation to the environment, and in this the explanation works as a complement rather that a rival to other views. It is clear from the paper's arguments how the proposal applies to such cases, explaining them in terms of interpersonal synergies, shared relevance of affordances, and improvement of grip.

2. aim at an encompassing account, that covers all different cases of skilled we-intentionality. If the authors were to follow the second route, they would need to have a different focus for the paper and make correlative adjustments in the main arguments. This might require that some of characterisations of joint action in terms of 'shared relevance' and 'improvement of grip' as applying to all cases, be dropped, for broadening the scope of these terms too much, risks turning them vague or empty. A good strategy here would be to analyse the different kinds of cases of joint action in more detail, explaining in what sense different cases of joint action can be accounted for with those same concepts, or nuances thereof, and in what sense any of these do involve any goals or intentions on the part of agents, that specify further aims, intentions, or goals that the agents aim to fulfil by acting in that way.

While both alternatives require refinements in the paper's claims, option (1) seems more feasible in the light of the materials discussed in the paper. Furthermore, as explained in more detail in Section 3, the arguments against rival theories of joint action are straightforward in this case, while they are not if the scope is broadened to cover all sorts of different cases. In this version of the paper, the aim of broadening the scope and applying the skilled intentionality tools to all cases of joint action could be mentioned but left for detailed discussion in a different paper (or several), entirely devoted to making those points.

3. Accounts of joint action have several different aims, some of which the account presented by the authors addresses, while others not.

It would be good for the authors to clarify what exactly their account aims to do with regards to these different debates and be explicit about it.

a) One of the main aims of theories of joint action is to explain what differentiates joint action and an action done by two individuals but in parallel.

The paper does not explicitly address this point. Yet, it might be that some ideas in the paper - e.g. interpersonal synergy - could account for this difference at least for cases of bodily coordination. However, if the paper aims to cover other cases, such as the 'painting the house' where joint action is remote in space and time, more argument would be required to support this claim.

b) Another aim that many views on joint action have is to position themselves on whether we can give an account of joint action in terms of individual states plus other conditions, or not. This amounts to asking whether 'jointness' in 'joint action' can be accounted for in terms that do not

presuppose jointness/sharedness or not. Reductionist accounts (e.g. Bratman, Ludwig) claim that we can, antireductionist accounts (e.g. Gilbert, Searle, Schmid) deny it.

The paper only briefly refers to this debate, in its discussion of Searle's view that mainly concerns, as with Bratman's, (c) below (i.e. whether inter-agential coordination can/should be explained in terms intentions). Now, one can wonder whether the 'Skilled-intentionality' view, in its own terms, is anti-reductionist with respect to its 'jointness'.

The position is not clear in the paper, and there are no arguments explicitly addressing this question. On the one hand, in as far as 'shared relevance' and 'improvement of grip' explain joint action in terms of reducing disattunment of the agents-environment systems, the view could be seen as reductionist. Yet, this changes when we consider the cases that require framing in terms of further shared activities, like the case of shopping for a home-renovation. In those cases, it seems that 'shared activities' – decorating the house – explain 'shared relevance' – looking at this particular tool in the DIY-store – and these shared activities in turn are explained by 'a history of interactions' – the couple plans to renovate their house, etc. In this line of argument, 'sharedness' might be seen as a primitive concept, for it leads to other forms of sharedness that together might be thought to form an explanatory circle.

Thus, here again we find an internal tension and an ambiguity in the view, for different cases seem to require different accounts. Even if one thought, as the paper also claims, that affordances are always nested in other affordances, and thus that all cases are like cases of the second kind, this requires further explaining. In any case, whether this is the authors' view is speculative, for this issue is not explicitly discussed in the paper.

c) Some views on joint action attempt to explain 'coordination in joint action' and discuss what elements are required to explain this, e.g. whether intentions or plans are required.

The paper gives an account of how coordination might work in many cases, relying on affordances provided by the living environment without the need of agents having plans and following up on them. In this, the paper is successful. In this respect, the proposal is indeed different from Bratman's, as the paper claims. Yet, it is important that the paper clarifies explicitly that these are not rival accounts, as Bratman states explicitly that his view is meant only as offering sufficient conditions – not necessary ones – for shared intentional activity, and the authors do not disagree that there might be cases of joint action where there are such explicit plans. In p.4, they explicitly say this: "The account of interlocking intentions Bratman provides is perhaps well-suited to account for the negotiations that might unfold between people when they plan to perform an action together (e.g., Bratman, 2014). However, it does not explain how multiple agents succeed in putting the plans they have made into effect, coordinating what they are doing when they act together in a particular situation." While this last point might be true, Bratman does not claim to be explaining this, for Bratmanian joint action might involve no bodily coordination at all. A Bratmanian case of joint action may be like the example of agents coming to paint the house at different times, mentioned above, where there is no face-to-face interaction, and agents have explicit plans and common knowledge of those. As long as there are some cases in which there are individual intentions and common plans, Bratman's account is not challenged.

In any case, beyond Bratman, the paper shows how intentions need not be considered necessary for inter-agential action coordination to occur. Yet the authors claim that they are offering an

account of joint action or joint activity that does away with intentions or goals altogether (see e.g. “Our aim [...]is to show how the work that shared goals and intentions are asked to do in philosophical accounts of joint action can instead be performed by affordances of shared relevance.”, p.5). This - I think - is not shown in the paper. While the arguments might show this is the case for spontaneous actions that involve bodily coordination between individuals, it is not clear from the considerations in the paper that this is so in all cases that the paper mentions.

d) Finally, some accounts are interested in arguing that the “sharedness” of joint action is dependent on the presence of a particular subject, a first-person plural subject or a ‘we’, where ‘we’ is understood as a plural agent not reducible to a mere sum of individuals (see Schmid and Gilbert for paradigmatic views).

The paper does not take a position on this aspect of the discussion. This is in part due to the fact that the paper does not clarify what is meant by ‘we’ in “we-intentionality”. (See (4) below for some remarks on this issue.)

In sum, in relation to the current scholarly discussion on joint action, the account presented in the paper clearly addresses (c) by making the case that affordances in the living environment, that are of shared relevance for the individuals, can account for inter-agential action coordination in the absence of intentions directed at common plans (or subplans that fit a shared plan). As said, the paper’s criticism of the literature on this issue is partly successful, that is, in as far as it focuses on the explanation of action coordination cases that do not require a further specification of the agents’ activities in terms of further joint aims, shared activities or shared goals of the agents. Yet it is not clear that the view succeeds as a general alternative to Searle’s or Bratman’s accounts when all cases of shared activity are included. As said above, this will require more and different kinds of argument.

Thus, while the paper could succeed in explaining coordination in action, it should clearly state its scope and aim as explaining how coordination is possible without intentions, rather than claiming that this is the case for all possible cases of shared activity, and attempting to cover a wide range of very different cases with the concepts of “shared relevance” and “improvement of grip”. In this regard, I recommend revision of the way the paper describes its aim, and correlative adjustments in the description of its conclusion, as well as examples, in line with (1) in (2) above. The paper would be best presented as mainly centred around (c), and addressing (a) and (b) in as far as they are corollaries of the analysis of the relevant cases, such as the double-handed saw. These corollaries should be drawn as explicitly addressing these questions for the relevant set of cases, i.e. those cases where an interpersonal synergy is developed as a way of improving grip for the agents involved.

In the case of (b), as said above, the authors might endorse a form of explanatory reductionism that explains joint action in terms of shared relevance, and that, in turn, in terms of improvement of grip. Alternatively, they may endorse a form of antireductionism, where ‘sharedness’ is seen as a primitive concept that leads to other forms of sharedness that together form an explanatory circle. It should be noted though, that to follow the second line of reasoning has consequences for the overall argument and aim of the paper. For according to such anti-reductionism, there is no explanation of ‘joint’ in joint action ((a) above), and a different answer to (c) follows as well. For, even if the account explains coordination in joint action, it is not obvious that an antireductionism of this kind does away with intentions altogether, since those might be part of the further

framework in which action coordination is nested, and thus agents' intentions might be part of what is required to explain how action coordination works *in toto*. Thus, for this kind of anti-reductionist view, it won't be (obviously) true that agents "do not need to have predefined intentions or goals that specify who is to do what and when, in order to coordinate with each other" even in the cases where interpersonal synergy builds from bodily coordination (e.g. the two-handed saw case). While in this respect the view might do away with intentions that explain *the steps* in the coordination involved in carrying out an activity (in this the paper might still complement Bratman's and Searle's accounts), at least in cases where interpersonal bodily synergies are involved, this is not to do away with any intentions/goals specifying the shared activity as such. These joint actions might be seen as embedded in other shared activities that involve more mediated outcomes, and are carried out by agents acting in virtue of complex shared values or norms, which in turn specify goals or intentions for the agents involved.

4 A final concern relates to the use of the term "we-intentionality" in the paper.

Not only do different authors mean quite different things by that term, but the paper does not address the key questions that surround that concept in the relevant literature. One set of issues relates to what is meant by 'we' in "we-intentionality", and another set of issues to what is meant by "intentionality". Both are a matter of controversy.

The authors characterise "we-intentionality" in the following terms: "what individuals do when they engage in group ways of thinking, feeling and acting" (see abstract, and cfr. p.3). Yet the focus of the paper is much narrower. It focuses on how the environment supports joint action and has much less or almost anything to say about shared feelings (there is a mention that affectivity is always involved in shared activity) or shared thoughts (which do not figure at all). The same applies to the reference to groups in the previous quote. The paper does not discuss in any detail how groups behave, beyond the case of dyads interacting. So, I recommend narrowing the focus to the main target of the paper, with an exclusive and explicit focus on joint action of dyads - and especially coordinating activities within those, perhaps within the larger topic of we-intentionality. The use of this term, if kept, will minimally require an explanatory footnote, noting some of the relevant discussions and some further references to others in the field using the term in related ways.

References

1. Butterfill S: Joint Action and Development. *The Philosophical Quarterly*. 2012; **62** (246): 23-47 Publisher Full Text
2. Butterfill S: Joint action: a minimalist approach. *The Routledge Handbook of Philosophy of the Social Mind*. 2016. Publisher Full Text
3. Fiebech A, Gallagher S: 'Joint attention in joint action'. *Philosophical Psychology*. 2013; **26** (4): 571-587 Publisher Full Text
4. Knoblich G, Butterfill S, Sebanz N: Psychological Research on Joint Action: Theory and Data. *Psychology of Learning and Motivation, Academic Press*. 2011; **54**: 59-101 Publisher Full Text
5. Satne G: Practical knowledge and shared agency: pluralizing the Anscombean view. *Inquiry*. 2020. 1-28 Publisher Full Text
6. Satne G: Understanding others by doing things together: an enactive account. *Synthese*. 2021; **198** (S1): 507-528 Publisher Full Text

Is the work original in terms of material and argument?

Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?
Yes

Is the work clearly and cogently presented?
Partly

Is the argument persuasive and supported by evidence?
Partly

If any, are all the source data and materials underlying the results available?
No source data required

Does the research article contribute to the cultural, historical, social understanding of the field?
Yes

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 08 Sep 2021

Julian Kiverstein

The aims of our paper *The authors need... to clarify their proposal and make it cogent. Depending on how they address these, they could follow two alternative paths:*

1. *circumscribe the focus of the paper to explain bodily coordination in joint action.*
2. *aim at an encompassing account, that covers all different cases of skilled we-intentionality. This might require that some of characterisations of joint action in terms of 'shared relevance' and 'improvement of grip' as applying to all cases, be dropped, for broadening the scope of these terms too much, risks turning them vague or empty.*

Author's reply: Our aim was to provide a framework for understanding we-intentionality – understood as what individuals do when they think, act and feel as a member of a group – *in terms of skilled intentionality*. We therefore resisted the more modest framing suggested by Prof Satne under option 1. We define 'skilled intentionality' as the selective responsiveness to multiple relevant affordances at the same time. The definition of we-intentionality we propose is intended to be maximally broad, covering not only cases of joint action but also group ways of thinking, acting and feeling that are often discussed under the heading of collective intentionality. Thus it would be underselling the ambitions of the paper to present our account as a theory of bodily coordination in joint action (option 1 above). To adopt such a framing for the paper would suggest that skilled intentionality is primarily a theory of perception and action rather than providing a conceptual framework

for understanding cognition across the board (so-called 'higher' cognition' (sapience) as well as 'lower cognition' (sentience)). We have added a few sentences to clarify the programmatic aims of our paper to apply the skilled intentionality framework to we-intentionality. We do not claim to have provided a fully worked out account of we-intentionality that can be applied to all possible cases. Our aim is to instead provide some useful conceptual tools that will enable this kind of work to be done by ourselves, and others in the future.

The explanatory limits of the Skilled Intentionality Framework *(In) many cases of shared activity, there is no clear 'grip' improvement that is taking place. For example, imagine a crowd in a stadium joining in a 'Mexican wave' to celebrate the victory of their soccer team and someone sitting in the same stand as these winning team supporters, but who is not a supporter of that team and does not join in the wave. Later, her own team scores a goal and she celebrates by singing a song that the supporters of her team are singing at the other side of the stadium. How should we think of her not joining the wave in terms of the view proposed? Another interesting kind of case is the following: I am about to cross the street and I see an old lady struggling to step down the sidewalk. I step in to help her. It is not clear what the affordances of shared relevance are in this case. The disanalogy here runs deep, for if we were to keep the same terminology to explain how individuals are motivated to act and coordinate in all these cases, the terms would become highly metaphorical. Can we explain Mexican waves or helping an old lady cross the street as agents' ways of improving their grip in the agent-environment system or do we need in addition, and crucially, reference to the agents' intentions, values or group memberships? These are cases in which socio-cultural norms and values in place partly explain what the agents are doing and why. In this, these cases are quite different than handling a piano together or a double-handed saw. In the examples I provided though, whereas in some cases there is bodily coordination (e.g., the stadium wave; helping an old lady cross the street), this is either not a straightforward case of improving grip, and/or the motivations of the agents are not completely explainable independently of further values, intentions and aims. For these kinds of cases, we cannot explain what the affordances of shared relevance are merely in terms of improvement of grip, where this is understood in terms of reduction of disattunment of the agent-environment system by the sole means of forming interpersonal synergies.*

Author's reply: We are grateful to Prof Satne for this challenge, and for presenting us with these intriguing apparent counterexamples. We think we can accommodate the cases she has described by better foregrounding the role that situated normativity plays in tending towards an improved grip. 'Situated normativity' is the term we have used in previous work to refer to the ability of skilled individuals to distinguish better from worse, adequate from inadequate, appropriate from inappropriate, or correct from incorrect in the context of a particular situation (Rietveld 2008). Improvements in grip are sensed by the individual based on situated normativity, also, importantly, in cases of so-called "higher" cognition or what we have called *reflective* situated normativity (van den Herik & Rietveld 2021). Reflective situated normativity is a form of normativity that relies on linguistic practices and involve explicit articulation of rules. If I tell my children: "No phones at the dinner table!", this is an example of reflective situated normativity. The articulated rule attracts attention to and puts on display a pattern of behaviour as inappropriate. It can therefore be characterised as reflective. Crucially the ability to make normative distinctions between better or worse, adequate or inadequate in acting skilfully is one that human agents develop and cultivate by taking part in social and cultural practices. Situated normativity is

always in play whenever the individual acts to improve grip on the field of relevant affordances. We would not agree that the notion of grip requires supplementing with "further values, intentions and aims" as Prof Satne suggests. We would instead propose that this explanatory work can instead be achieved by reflecting on the role of situated normativity in regulating the agent-environment dynamics such that the agent tends towards improved grip in acting skilfully. In our revisions, we show how this might work in the case of helping the old lady to cross the street (in footnote 12, p.17). Grip increases by helping her because one would experience what we call 'discontent' (an affective tension) were one to leave her to make her way across on her own. Acting to help her thereby, on the other hand, reduces the feeling of discontent that is experienced as long as she appears to be at risk. A similar line of argument could be made for the example of the Mexican wave, though we did not add this example to the paper. The person that does not participate in the Mexican wave refrains because her taking part would feel wrong. The invitation to join in from others elicits tension, which she reduces by remaining in her seat. Thus, her declining to take part is a way for her to improve grip.

The explanatory role of predefined goals and intentions *In the case of helping the old lady, or in their own examples of "hunting for berries" or "shopping for a home renovation"... it seems that we need further explanation of what the agents are doing and those involve reference to something more: some goals, intentions, values, or social norms that the agents share. 'Shopping for a home-renovation', 'being in the silent carriage', 'painting a house together at different times', they all involve some pre-determined goals, intentions or norms in terms of which who takes part in a joint action and when are determined. Even if the authors would want to resist this claim, much more argument would be needed.*

Author's reply: As already noted, we fully agree that normativity is crucial and that is why we coined (and extensively argued for) the notions of unreflective situated normativity (Rietveld 2008) and more recently reflective situated normativity (van den Herik & Rietveld 2021), see our previous reply. However, as we note on p.10 of our paper, we agree with Lucy Suchman that the explanatory appeal to pre-determined goals or intentions, which is very common in the philosophy of cognitive science, explains very little. How an agent constrains and directs their action in the light of this goal or intention remains something of a mystery. We suggest the explanatory work of these constructs can be taken over instead by the self-organising dynamics of the agent-environment system. We have developed the Skilled Intentionality Framework as a conceptual framework for a post-cognitivist cognitive science that aims to take concepts such as goal and intention and provide tools for making sense of these concepts in the terms of ecological-enactive cognitive science. Now clearly it would be a problem for us if the conceptual tools we were proposing somehow had to rely on goals and intentions to do their explanatory work. We would face an objection of tacitly appealing to the very concepts we seek to render intelligible. However, so far as we can tell our account is not vulnerable to such an objection. For instance, we do not need to appeal to plans, intentions or goals to explain why the couple went shopping for DIY materials. We instead appeal to what we call 'large-scale affordances' of renovating the house - small-scale affordances such as all of the activities that contribute to our visiting the DIY shop for materials weave together over time because of our responsiveness to the large scale affordance of renovating the house (see also our ethnography of a shopping trip to a carpet warehouse in van Dijk & Rietveld, 2018: p.12-13). Do we not need to appeal to the goal of

renovating their house to explain why the couple undertake such a shopping trip in the first place? Our proposal is to have the work of this overarching plan taken over by tending towards an improved grip on a large-scale affordance. We connect this self-organising dynamic to free-energy minimisation in other work (see e.g. Bruineberg & Rietveld 2019). Thus, we are able to provide some empirical support for this proposal based on state-of-the-art work in theoretical neurobiology. SIF has the explanatory resources to unpack how to make sense of such a self-organising process of future-directed action without having recourse to plans, goals and intentions.

Doing justice to the different projects in the literature on we-intentionality *It would be good for the authors to clarify what exactly their account aims to do with regards to these different debates and be explicit about it.*

1. *what differentiates joint action and an action done by two individuals but in parallel?*

Author's reply: This question is addressed in the section headed Interpersonal synergies. Interpersonal synergies make possible a pooling of individual abilities that make joint actions possible that would not be possible for individuals acting in parallel.

1. *Can 'jointness' in 'joint action' be accounted for in terms that do not presuppose jointness/sharedness or not?*

Author's reply: Our aim is to have (1) affordances of shared relevance and (2) correlative abilities acquired by taking part in sociomaterial practices take over the explanatory work that interlocking intentions are asked to do in, for example, Bratman's account of action coordination. The sharing of affordances can be understood in two ways in our account. First, the landscape of affordances is defined in relation to a *shared* form of life (and that can be a sociomaterial practice in the case of human beings). The affordances that the landscape makes available have formed through patterns of practice that are constituted by the coordinated behaviour of multiple individuals. What groups, dyads and individuals share is that they are situated in the rich landscape of affordances. Moreover, the affordances the landscape makes available owe their continued existence to the actions of individuals that contribute to the continued availability of those affordances. We see this in the example of the silent carriage where the individual contributes to maintaining a sociocultural practice. Second, we account for joint action in terms of how interpersonal synergies form between individuals. Affordance related states of action readiness form in each of the individuals as they tend towards an improved grip on the environment (Bruineberg & Rietveld, 2019). Action readiness is a state of the whole body that is simultaneously behavioural and affective (Frijda 2007). What is different in cases of joint action is that the abilities that account for the formation of patterns of action readiness in individuals are pooled which has the consequence that affordances can stand out from the landscape as being of *shared* relevance.

1. *Some views on joint action attempt to explain 'coordination in joint action' and discuss what elements are required to explain this, e.g. whether intentions or plans are required.*

Author's reply: We addressed this objection above in our response to your first query about the scope of our argument. The aim of our framework is to account for the explanatory work that traditionally intentions and plans are supposed to do in terms of selective and simultaneous responsiveness to multiple relevant affordances, including large scale affordances.

1. *some accounts are interested in arguing that the "sharedness" of joint action is dependent*

on the presence of a particular subject, a first-person plural subject or a 'we', where 'we' is understood as a plural agent not reducible to a mere sum of individuals (see Schmid and Gilbert for paradigmatic views).

Author's reply: We do not directly address this question in our paper, but we have now added a footnote to explain the form that this question takes in our account. Something like this question comes when we consider whether the pooling of abilities that happens when interpersonal synergies form results in a group agent or not. If we say that individuals that are acting as parts of an interpersonal synergy form a group agent, this would imply an important difference between skilled intentionality and skilled *we*-intentionality. In the former the skilled agent is an individual, while in skilled we-intentionality the agent is a dyad or a larger group. We claim however that the difference is best understood in terms of timescales, and thus there is no need for us to invoke a group agent. Over long timescales, individuals are always coordinating with the multiple members of a sociomaterial practice. We saw this in the example of the silent train carriage. Over shorter time scales, this coordination requires the pooling of abilities in ways that lead to the formation of interpersonal synergies.

References Bratman, M.E. 1993. Shared intention. *Ethics* 104: 97-113. Bruineberg, J. & Rietveld, E. 2019. What's inside your head once you've figured out what your head is inside of. *Ecological Psychology* 31 (3): 198-217. Frijda, N.H., 2007. *The Laws of Emotion*, Mahwah, NJ: Lawrence Erlbaum Associates, Inc. Gallagher, S. & Zahavi, D. 2008. *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science.* (London and New York, NY: Routledge, 2008). Heft, H. 2001. *Ecological Psychology in Context: James Gibson, Roger Barker, and the Legacy of William James' Radical Empiricism.* New York: Taylor & Francis/Psychology Press.. Kaufer, S. & Chemero, A. 2015. *Phenomenology: An Introduction.* Cambridge, UK: Polity Press. Knoblich, G., Butterfill, S., & Sebanz, N. 2011. Psychological research on joint action: theory and data. In B. Ross (Ed.) *The Psychology of Learning and Motivation* Vol. 54. Burlington MA,: Elsevier Academic Press, pp.59-101. Reed, E. 1996. *Encountering the world: Toward an ecological psychology*. New York, NY: Oxford University Press. Richardson, M.J., Marsh, K.L., & Baron, R.M. 2007. Judging and actualising intrapersonal and interpersonal affordances. *Journal of Experimental Psychology: Human Perception and Performance.* 33: 845-859. Rietveld, E., & Kiverstein, J. (2014). A rich landscape of affordances. *Ecological Psychology*, *26*(4), 325–352. doi:10.1080/10407413.2014.958035. Rietveld, E. (2008). Situated Normativity: The Normative Aspect of Embodied Cognition in Unreflective Action. *Mind*. Tollefsen, D. & Dale, R. 2011. Naturalising joint action: a process-based approach. *Philosophical Psychology* 25 (3): 385-407. Tomasello, M. 2019. *Becoming Human: A Theory of Ontogeny*. Cambridge, MA: Harvard University Press. Van Den Herik, J. & Rietveld, E. 2020. Reflective situated normativity. *Philosophical Studies* Van Dijk, L. & Rietveld, E. 2017. Foregrounding sociomaterial practice in our understanding of affordances: the skilled intentionality framework.

Competing Interests: No competing interests were disclosed.