

Situating Joint Action

Abstract: Whether action is joint and conforms to agents' commitments, and what capacities agents must exercise in order to engage in joint action, is a matter of the social structure of the situation of interaction, not of the agents' believed sharing of mental states.

KEYWORDS: Joint action, Shared expectations, Status functions, Deontology.

1. Introduction

People act together – we go for walks, cook, sport, vote, etc., together. People personate roles and functions when acting together – we act as friends, colleagues, in authority positions, as subordinates, etc. People, when acting together, are situated in contexts – we act in the marketplace, around dinner tables, in church, at work, etc. People acting together in contexts where they personate roles and functions are committed to act in agreement with certain expectations – e.g., policemen patrolling are expected to maintain order and teachers when holding class are expected to follow the curriculum.

The concept 'joint action' is meant to pick out instances of acting together where at least two agents contribute in complementary ways with part performances to what they do together. The philosophically interesting question is: Under what conditions do a set of part actions qualify as a joint action? It is important to distinguish this *conceptual* question from the *genealogical* questions: How do joint actions arise? And relatedly: How are agents able to engage in joint action? Answering the first question is to demarcate the class of joint actions by stating necessary and/or sufficient conditions for an instance of acting together to qualify as a member of that class, while answering the latter is to specify cognitive processes and/or capacities employed by agents when engaging in joint actions.

The slide between conceptual and genealogical focus is obvious in today's theorising on joint action. There are roughly two camps of theories of joint action. Elizabeth Pacherie (2011) has called them 'maximalism' and 'minimalism'. On the maximalist side we find philosophers such as Michael Bratman, Margaret Gilbert, and Raimo Tuomela who argue that for a set of part actions to qualify as a joint action it is necessary that participants' mental states be appropriately related. Specifically, participants must have common knowledge or mutual belief about the mental component – intentions, beliefs, or commitments – of the activity. On the minimalist side we find developmental psychologists and neuroscientists such as Celia Brownell, Günther Knoblich, and Natalie Sebanz who propose that joint

Patrizio Lo Presti
Department of Philosophy and Cognitive Science
Kungshuset 222 22, Lund University, Lund, Sweden
e-mail (✉): Patrizio.Lo_Presti@fil.lu.se

attention, action observation, and shared goal representations suffice *for agents to be able to act jointly*. Minimalism is informed by data in neuroscience (Knoblich and Jordan 2002, 2003) and developmental psychology (Tomasello et. al. 2005; Rakoczy 2006; Tomasello and Rakoczy 2007). Data suggests, on the one hand, that action observation trigger shared goal representations that allow joint action, and, on the other hand, that children engage in joint action, presumably prior to development of capacities for having common knowledge or mutual belief about the mental component of the activity as stated as necessary for joint action according to maximalists.

Apparently, the dispute between maximalism and minimalism as these camps are here depicted is about *methodological* and about scientific priority. If, for example, children perform kinds of actions recounted in the first paragraph in pretence play, this can be taken to indicate a *genealogical* fact about joint actions that forces revision of the maximalist *conception* of joint action. But, it can also be taken to indicate that although children act together, their actions are not properly *joint* actions. The issue is thus methodological: should philosophical analysis be attentive to accommodate experimental data or should empirical scientists be bound by philosophical analysis when interpreting data? In other words, is it conceptualisation or observed genealogy that sets the agenda for whether an instance of acting together is joint activity? Either way of prioritising has its costs. Too general conceptions blind us to important distinctions. In the case of joint action there is risk of conflating any instance of coordination in traffic or crowds with joint projects of groups. But then again, too conservative conceptions blind us to the amplitude of instances of phenomena. In the case of joint action there is risk that the class of joint actions is diminishingly small and relegated to groups kept together by shared higher order thoughts.

Middle ground is breaking. On the one hand, minimalists are qualifying their claims about joint action. Sebanz et. al. (2006) and Vesper et. al. (2010) speak of *simple kind* joint actions, Brownell (2011) accounts for joint action *early in development*, Michael (2011) speak of *emotionally guided* joint action. On the other hand, maximalists are conceding that they are focusing specific kinds of joint action. Tuomela's distinctions between levels of togetherness (2007) are perhaps the hallmark of recognition of the versatility of joint actions. Pettit and Schweikard (2006) distinguishes group agency from joint agency, and Pettit and List (2011) speak of corporate agency.

It seems we have arrived, then, at a fair methodological compromise preserving both conceptual distinctiveness and recognition of the amplitude of the class of joint actions as indicated by genealogical facts found by developmental psychologists and neuroscientists. We now stand on common ground, surrounded with hoards of kinds of 'joint action' to study.

2. Situating joint action

I think the study of joint action is in dire need of complementary perspective. We have seen the positive differences between maximalist and minimalist analyses of joint action. Now I want to focus their negative similarities.

Similar to maximalism and minimalism is failure to appreciate that agents acting together are situated in contexts where they personate different roles and functions.

Consider the examples of acting together in the introduction. Note how, for each example of acting together, if we alter some *feature in context* of interaction the situation suddenly looks very different. First, for each example given we can introduce *statuses, functions, and roles* of participants, such as police officers, colleagues, friends, priests, and so on. Second, for each example we can think of *different arenas* in which people act together, such as in the police station, at work, at home, in church, and so on. Third, imagine any of the examples of acting together in some arena, now shift between introducing and removing some object in the imagined situation *indicating the roles or statuses of persons involved*. Take this simple example: You are out driving. The car behind you starts flashing its headlights. You wonder what the other intends, what the other believes that you intend, what your commitments are, if any. Perhaps you veer to let the other pass. That is a fair guess of what the other wanted you to do. Now imagine the same situation but introduce an indicator of some role or function, e.g., car lacquer or sirens, indicating the car to be an ambulance. Perhaps you veer to let the other pass because, after all, that is what you are supposed to do in this situation. Now ask yourself: What did you need to know or believe about the other's mental states in these situations for you to take appropriate action? If your intuitions differ it may have to do with how the situations differ. In the latter situation but not the former the other personated a certain role, 'ambulance driver'. In the latter situation but not the former the other's role was indicated by something to which you attended, sirens or car lacquer. Intuitively, a feature of context that you attended to non-coincidentally facilitated coordination of actions in the latter situation.

Features of context, I will argue, short circuits necessity of shared higher order thoughts, as stated as necessary in one of the condition in maximalist analyses of joint action, while demanding only lower levels of attention (Brinck, Zlatev, and Andrén 2006) than joint attention and action observation, as conjectured by some minimalists. The claim I am interested in is that *there are instances of joint action where context of interaction obsoletes the necessity of common knowledge or mutual beliefs about mental states indiscriminately of kind of joint action*. If I am right about this, then there are reasons to count as joint actions actions not covered by the received philosophical view. Importantly, if I am right, the common ground at which we stand as philosophers and empirical scientists, where we find sets of different kinds of joint actions can be put in new light reflected by all those kinds. It is in the light of the structure of social reality in general and the structure of

contexts of interaction specifically that we have to view the phenomenon of joint action. The following section explains the details.

3. Social structure of contexts of interaction

Features in context of interaction indicate roles and functions of persons and objects in the context. Functions and roles entail procedures of interaction that are permitted, obliged, forbidden, etc., in context of interaction. To explain this I will give a short recount of a prominent theory of the structure of social reality.

John Searle's ontology of social reality (1995, 2003, 2006, 2010) has enjoyed abundance of attention from philosophers and developmental psychologists. Searle's aim is, in short, to provide a theory of how there can be an *ontologically subjective but epistemically objective social reality* of money, marriage, nation states, and so on, on top of the *ontologically objective physical reality* of molecular bonds of pieces of paper, mountains, and particles in fields of force. There are four points essential to understanding Searle's social ontology and how it is related to my claim.

First we have *status functions*. Status functions are roles of persons and object in social environs, what they 'count as'. For instance, some people count as police officers, others as teachers, and others as married couples. The now famous 'X counts as Y in C' formula is shorthand for saying that something or someone *X* has a certain role or function *Y*, e.g. police officer, teacher, etc., in some context *C*, e.g., country, workplace, meeting, etc.

Second we need to know how status functions are *created*. Status functions are created by status function declaration (SFD). SFDs are utterances the propositional content of which represent a state of affairs where the subject (*X*), is declared to have a function (*Y*) in a context (*C*) – for instance, "Ban Ki-moon (*X*) counts as Secretary General (*Y*) of the U.N. (*C*)". For a SFD to be *successful*, that is, for the state of affairs represented to apply, it must both be *collectively recognised*, and it must be *in virtue of that recognition* that the *Y* status apply to the object/subject *X* in the context *C*. Collective recognition is attitude of multitude of agents, "We recognise, X counts as Y in C". Now, for status functions to *persist* they must be continuously collectively recognised as applying and people must act accordingly. For instance, people must continuously recognise certain pieces of paper to count as medium of exchange and to treat them as such.

Third, status functions entail *deontology*. Deontology is range of permitted, prohibited, obliged, etc., actions or uses of *X* by virtue its *Y* status in context of action *C*. For instance, counting as Secretary General of the U.N. entails certain rights and obligations to hold summits, address the General Assembly, and so on, within the context of the charter of the U.N.

Fourth, status functions have *indicators*. Indicators are objects assigned the function to indicate what status functions apply in contexts of interaction. For instance, we have police badges, wedding rings, identity cards, and so on, counting

as indicative of people counting as police officers, married, citizens, and so on. Indicators of status functions allow people to identify what status functions apply in contexts in which they interact. Owners of property have receipts or contracts, which can function as proof if the status of ownership is questioned, voters verify status as citizen with identity cards (where voting is part of the deontology entailed by the status function ‘citizen’), ambulances are indicated by sirens and car lacquer – and so on for most instances of ‘X counts as Y in C’.

It is important now to relate these four points – status functions, creation and persistence of status functions, deontology, and indicators – to contexts of interaction generally and instances of acting together specifically. The first relation is easy to see. Think of any interaction between two or more agents. Necessarily, the agents will be situated in some specific context. Each agent will likely personate some role recognised by the other in that context. By virtue of these roles it is likely that there are action procedures in interaction that the agents are obliged to, permitted to, or prohibited from carrying out. It is also likely that there are objects in the context indicating the roles that apply. By implication from status functions to deontology, indicators are indicative of what action procedures are obligatory, permitted, or prohibited in the context of interaction.

Importantly, agents in contexts of interaction structured in the Searlean sense will be in position to *expect* each other to act in certain ways. How so? Collective recognition of status functions is necessary condition for existence and persistence of status functions, hence also for deontology entailed by status functions. Indicators of status functions, therefore, are indicators of a range of action expected in context of interaction – the range of actions agents are committed to by virtue of the status functions that apply to them. Now, the central point is this: *interaction is situated activity in contexts structured around roles and functions allowing, permitting, or prohibiting certain action procedures indicated by objects in the particular situation*. Thus, attention to indicators in context inform agents about action procedures expected in context without their attending to, or having doxastic states the content of which refer to, any *specific* other agent in the context in which the agents are situated.

Before I turn to *joint* action, remember that according to maximalist analyses of joint action it is a necessary condition for part actions to qualify as a joint action that the parts are performed in common knowledge or mutual beliefs about intentions and/or commitments to action procedures (at least for non-simple action kinds). Remember also that, according to minimalist accounts of joint action, it is sufficient that agents, when performing their parts, engage in joint attention and action observation, which excite shared goal representations (at least for simple kind joint actions). Consider this suggestion – a recollection of the ambulance example in the preceding paragraph: part actions performed in accordance with expectations about action procedure triggered by attention to indicators of status functions allow non-coincidental coordination of part actions into joint action.

3. Implications of the situated approach to joint action

The implications of my suggestion are manifold. First, we can count as joint action any set of actions non-coincidentally performed in conformity with collectively recognised deontology. Sets of part actions performed in conformity with deontology will be non-coincidentally performed if performed out of expectations triggered by features of context indicating roles and functions and thus also indicating action procedures. Second, as a consequence, we should revise conceptions of joint action according to which common knowledge or mutual belief among participants about each others' doxastic or motivational states is necessary for joint action. This second implication is not based on observed genealogical fact, but on an analysis of the structure of social reality in which agents are situated when acting together. Thus, there are conceptual reasons for revising the maximalist conception of joint action. Third, the conception of joint action that emerge is that lower level of attention than joint attention and action observation suffices for sets of part actions to qualify as joint action. It suffices that agents are attentive to features of context indicating appropriate action procedures in that context. That is, there need be neither triangulation of attentional states between agents and a shared object of attention, nor attention to each other among agents, for them to be able, genealogically speaking, to contribute non-coincidentally with complementary part performances to what they do together. Fourth, since both simple- and non-simple kinds of joint activity are situated activities, these implications are implications for simple and non-simple kinds of joint action alike. Attention to features of context suffices for joint action in highly institutionalised contexts of interaction if status functions, and hence deontology, are indicated by objects attended to. This is true also for simple-kind joint actions where, e.g., the object to obtain or state of affairs to alter that is the goal of the agents is perceptually accessible in the context of interaction (cf. Brinck and Gärdenfors 2003; Sebanz et. al. 2006; Vesper et. al. 2010). Fifth, my suggestion implies that the *conceptual* analyses of 'joint action' of philosophers and the *genealogical* accounts of psychologists and neuroscientists do not evoke methodological issues about analytical contra experimental priority. Minimalists have non-minimalist reasons for criticising the maximalist conception. Taking *situatedness* of joint action seriously now motivates criticism of maximalism. Sixth, taking situatedness of joint action seriously allow understanding of how all, or at least most, of the kinds of joint actions differentiated in the aftermath of philosophical-psychological issues of methodological priority are unified because tokened in social environs where people have roles and functions and stand in different relations which are indicated by features in those environs. To not take situation seriously when analysing joint action is, therefore, to miss out an aspect that permeates almost all instances of joint action, whether 'simple' or not. Seventh, of course, if there are *no* status functions or indicators thereof in some context of interaction, and when status functions are *first* jointly created, we might

have to resort to maximalism or minimalism, *or* the situated framework can be expanded to embrace also those instances of acting together. For example, Searle's theory of creation of status functions can be abandoned to the benefit of having a wider conception of roles, statuses, and functions and thus to the benefit of developing a situated framework autonomous from maximalism and minimalism.

4. Conclusion

Joint action is not genealogically restricted to actions contributed by people in complementary ways to reach a shared goal, with common knowledge or mutual beliefs about each others' mental or attentional states gluing the actions together. Joint action is not conceptually restricted to sets of actions performed in common knowledge or mutual belief. Joint action includes actions performed by agents sharing situation, where recognition of appropriate action procedures in the situation of interaction through attention to features of context glue the actions together.

References

- Bratman, M. (1999). *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge: Cambridge University Press.
- Brinck, I. and P. Gärdenfors (2003). 'Co-operation and Communication in Apes and Humans'. *Mind & Language* 18(5): 481-501.
- Brinck I., Zlatev, J. and M. Andrén (2006). 'An applied analysis of attentional intersubjectivity'. SEDSU Deliverable 7.
- Brownell, C. A. (2011). 'Early Developments in Joint Action'. *Review of Philosophy and Psychology* 2: 193-211.
- Gilbert, M. (1989). *On Social Facts*. Princeton: Princeton University Press.
- Gilbert, M. (2009). 'Shared intention and personal intentions'. *Philosophical Studies* 144: 167-187.
- Pettit, P. and C. List (2011). *Group Agency: the possibility, design, and status of corporate agents*. New York: Oxford University Press.
- Pettit, P. and D. Schweikard (2006). 'Joint Actions and Group Agents'. *Philosophy of the Social Sciences* 36(1): 18-39.
- Rakoczy, H. (2006). 'Pretend play and the development of collective intentionality'. *Cognitive Systems Research* 7: 113-127.
- Rakoczy, H. and M. Tomasello (2007). 'The Ontogeny of Social Ontology: Steps to Shared Intentionality and Status Functions'. In *Intentional Acts and Institutional Facts: Essays on John Searle's Social Ontology*, ed. Tsohatzidis, S. L. Dordrecht: Springer.
- Searle, J. (1995). *The Construction of Social Reality*. New York: The Free Press.
- Searle, J. (2003). 'Social Ontology and Political Power'. In *Socializing Metaphysics: The Nature of Social Reality*, ed. Schmitt, F. F. Lanham, Maryland: Rowman and Littlefield Publishers, Inc.
- Searle, J. (2006). 'Social ontology: Some basic principles'. *Anthropological Theory* 6(1): 12-29.
- Searle, J. (2010). *Making the Social World: The Structure of Human Civilization*. New York: Oxford University Press.

- Sebanz, N., Bekkering H. and G. Knoblich (2006). 'Joint action: bodies and minds moving together'. *TRENDS in Cognitive Science* 10(2): 70-76.
- Tuomela, R. (2007). *The Philosophy of Sociality: The Shared Point of View*. New York: Oxford University Press.
- Vesper, C., Butterfill, S., Knoblich, G. and N. Sebanz (2010). 'A minimal architecture for joint action'. *Neural Networks* 23: 998-1003.