Editorial Manager(tm) for The Review of Philosophy and Psychology
Manuscript Draft

Manuscript Number: ROPP106

Title: A case for weak joint action

Article Type: Joint Action: What is Shared?

Keywords: Joint action; common knowledge; group identification; shared attitudes; collective goals

Corresponding Author: Cedric Paternotte, Pd.D.

Corresponding Author's Institution: University of Bristol

First Author: Cedric Paternotte, Pd.D.

Order of Authors: Cedric Paternotte, Pd.D.

Abstract: Many definitions of joint action have been given over the last three decades. But for all their differences, they still have much in common. In particular and as a negative consequence, they cannot fit some perplexing cases of weak joint action, such as demonstrations, where agents' rely on distinct epistemic sources and as a result have no first hand knowledge about each other. I argue that such collective actions, typically anonymous one-shot interactions, can still receive an analysis that does not depart too much from the classical ones. To that end, the concept of common knowledge must be defended against misguided criticisms first, then adequately weakened. However, although the common knowledge thus obtained is weaker than the classical, in weak joint actions its content is rich. As a result, even if the links between individuals are seriously stretched in weak joint actions, much is still shared among them. Weak joint actions still require considerable cognitive abilities indeed.

Title page
Click here to download Title page: RPP_joint-action_title.doc

Name : Cédric Paternotte

Title : A case for disjoint action

Affiliation: University of Bristol, Department of philosophy

Address: 43 Woodland Road, Bristol BS8 1UU, U.K.

Tel: +44 (0) 117 33 18395

Email: cedric.paternotte@bristol.ac.uk

Acknowledegements : This work was supported by the Arts and Humanities Research Council of the UK, Grant No. AH/F017502/1, which I gratefully acknowledge.

MS
Click here to download Manuscript (without any authors' affiliations): WJA.pdf

A case for weak joint action

Abstract.

Many definitions of joint action have been given over the last three decades. But for all their differences, they still have much in common. In particular and as a negative consequence, they cannot fit some perplexing cases of weak joint action, such as demonstrations, where agents' rely on distinct epistemic sources and as a result have no first-hand knowledge about each other. I argue that such collective actions, typically anonymous one-shot interactions, can still receive an analysis that does not depart too much from the classical ones. To that end, the concept of common knowledge must be defended against misguided criticisms first, then adequately weakened. However, although the common knowledge thus obtained is weaker than the classical, in weak joint actions its content is rich. As a result, even if the links between individuals are seriously stretched in weak joint actions, much is still shared among them. Weak joint actions still require considerable cognitive abilities indeed.

1. Introduction

Consider individuals singing a duet together, or painting a house together, each of them choosing her actions in accordance with those of others. Or two persons walking together, adjusting their pace to one another's, expecting the other to do the same and disposed to react if she does not. Those are familiar examples of joint actions, or at least of what philosophers have considered as characteristic joint actions. The general idea is that in joint action, there exists some kind of link between the individuals, the nature of which should be elucidated. Combined actions of individuals can regularly or reliably produce a collective outcome because something is shared between them. But what exactly is shared? Depending on authors, the main aspect of such a link can be epistemic, deontic, intentional, teleological, and/or even strategic. The literature on joint actions is fraught with collective intentions, common beliefs, collective goals, group commitments, etc., some of which can be intertwined in more or less complex fashions. Individuals can have beliefs about the intentions, beliefs and goals of others; beliefs, goals and intentions can be reasons or causes for other beliefs, goals and intentions; and so on. As a result, it is sometimes hard to know where really lies the collective dimension of joint action, what makes it "joint" rather than a mere juxtaposition of individual actions.

© 2011 *Kluwer Academic Publishers. Printed in the Netherlands.*

However, it seems that the task can be achieved in most cases. Even if authors disagree, they all manage to identify some core of jointness, some elementary brick that grants an action its social dimension and allows one to identify what is shared between individuals. This is partly because these authors usually base their analysis on some intuitive, obvious cases of joint action. As a consequence, if they can identify what constitutes such actions, they give no guarantee that they have found the conceptual core of jointness.

Why would it be a problem in the first place? Because there exist some cases of joint action in which none of these classical requirements is met. In some cases of demonstrations or flash mobs, individuals lack knowledge about each other, and as a result, as we will see, such examples fail to meet the usual conditions for joint action. But intuitively, these are still kinds of joint action, even if extremely weak ones. People regularly participate in order to produce collective outcomes without knowing beforehand who else will participate and seem to do it intentionally and after deliberation rather than automatically or by habit. Can we do justice to this intuition that at least some of the required ingredients used to define joint action are still present in such cases?

The aim of this paper is to defend the seemingly paradoxical idea that weak but sophisticated forms of joint action can exist even if the epistemic structure between individuals is extremely loose, thanks to their complex epistemic content. First, section 2 briefly introduces classical analyses of joint action and highlights some of their common features, before arguing that they are too stringent to cover cases of weak joint actions. The following sections then turn to concepts that help understand such cases. In this paper's main section (section 3), I argue that any such definition must still use the concept of common knowledge, lest no definition can be said to concern joint actions anymore. Yet, in the analysis of weak joint actions, classical common knowledge should be weakened, not because it is too cognitively demanding (as it is commonly – and falsely – thought), but in order to correspond to the dimmer epistemic link that exists between individuals. I comment several accounts or criticisms of common knowledge, most of which are misguided in my view. Section 4 then sketches the psychological mechanisms that could underlie weak joint action, works out their implications as to what is shared between individuals, and draws some more general consequences of such an account of weak joint action. Overall, one aim of the paper is thus to get closer to a psychologically realistic analysis of weak joint actions. Note that this paper does *not* provide an analytical definition of weak joint action, which would necessitate a longer, fuller analysis. Rather, presupposing

that such a definition can be given, it discusses its epistemic aspects, namely its epistemic structure and content.[1]

2. Joint Action

I now introduce four classical accounts of joint action. Detailed analyses of them already abound in the existing literature, and I will be content to merely recall what each of them takes to be the core of joint action. It will then allow me to argue that none of them corresponds to cases of weak joint action, which is enough for the purposes of this paper.

2.1. Classical accounts

Bratman (1992, 1993) takes the previous examples of singing a duet or painting a house to be typical joint actions. In his vocabulary, agents have a shared intention to do J[2] whenever they all have individual intentions that they do J,[3] these intentions are caused and/or justified by meshing subplans of action, and all this is common knowledge between them. For now, let it suffice to say that it is equivalent to declare that something is common knowledge or that it is public (common knowledge will be more thoroughly analysed in section 3). As in most accounts, as we will see, joint intention is here constituted of individual intentions constrained by specific conditions, this being common knowledge; what varies among authors is the nature of these conditions. In Bratman's case, they stipulate the existence of meshing subplans of action. For agents to have meshing subplans does not mean

[1] Another presupposition of this paper is that joint action can generally be analysed in terms of sets of sufficient and necessary conditions, most of which only refer to characteristics of individuals rather than of groups. In particular, I will not consider purely holistic views akin to that of Searle (1990, 1995), who considers that joint intentions are irreducibly collective. Such views represent but a minor strand of the literature, and it seems to me that their acceptance entails an a priori rejection of any analytic definition of joint action. In other words, I will not discuss the general relevance of the search for such definitions. Pacherie (2003) discusses and criticizes such holistic views. The accounts discussed in section 2 are at best partially holistic, but all leave ample room to individualistic components; I adopt a similar stance.

[2] Authors alternatively try to define joint intentions or joint actions. The task is equivalent in both cases: a joint action is only what results from a joint intention, that is, nothing else than the set of actions caused by the joint intention of the set of individuals.

[3] These are individual intentions with a collective content, which as noticed before can be problematic if one wishes to endorse methodological individualism. This is also the main feature of Searle's holistic approach. I will stay silent on whether this feature is realistic or not, since I don't see it as central to Bratman's account and it has no bearing on my analysis.

that they need to have planned in advance the exact actions they will accomplish, but merely that they intend to act together in consistent ways, that is, by choosing actions that whenever possible do not prevent the other from accomplishing his part. In other words, the constraint Bratman puts on individual actions concerns their coherence, which agents should strive to achieve. When a joint action is only composed of sets of determined individual action (and that agents realise they are coherent), the coherence condition is automatically met. Similarly, it will be easily met when only a small set of individual actions can constitute the joint action. In particular, this will often be the case when a joint action is instantaneous, that is, when agents only have to choose one action each, the effect of which is limited. Only when joint action persists over time and is composed of sets of successive individual actions does the coherence condition really starts to constrain it. We can thus say that Bratman's constraints are mostly *diachronic*.

Gilbert (1989) defines a joint (or shared) action as an action resulting from what she calls "joint readiness", which is a situation where it is common knowledge among agents that they have mutually expressed their conditional intention to do their part if others do theirs.[4] This seems even simpler than Bratman's account: here we only have a set of conditional individual intentions that are common knowledge. However, the matter is complicated by Gilbert's repeated emphasis on the fact that such situations give rise to commitments of the agents to do their part.[5] This concept of commitment is not explicit in Gilbert's definition of shared intention or action, which makes it difficult to esteem whether being committed to do something is equivalent to having decided to do it, or if it corresponds to an obligation.[6] What matters is that Gilbert considers this aspect to be necessary to joint action, so that it could allow one to detect it: if a joint action was being accomplished, then when an agent stops doing her part, others feel entitled to react, and she should expect such a reaction (regardless of whether it happens or not). In other words, Gilbert's account emphasizes the *normative* aspect of joint actions.

Miller's tack on joint action highlights another crucial ingredient: collective goals.[7] He defines a joint action as a set of intentional individ-

[4] Taken from (Gilbert, 1989), p. 197. Gilbert calls this conditional intention "quasi-readiness", but the exact terminology does not matter here.

[5] More precisely to a joint commitment of the group, which is not an individualistic concept.

[6] This point was raised by Miller (2001), p. 85.

[7] Miller was not the first author to put collective goals at the heart of joint action; Tuomela had been doing so from the late 1980. I mention Miller's account for its simplicity, and because it involves no striking element apart from collective goals.

ual actions, each of which is accomplished in order to realise a collective end, all of this being common knowledge (Miller, 2001, p. 57). This time, it is the concept of collective end that lacks a precise definition. This is basically a set of individual ends that are all necessarily realised by an identical state of affairs: one could not be realised independently of the others (Miller, 2001, p. 58). Think for instance of a common benefit that cannot only be obtained if all cooperate. The important idea is that the core of the jointness lies in the collective end, rather than in intentions with a collective content or in joint commitment. The individual intentions must mutually justify themselves, but they must also stem from an identical reason to act. Accordingly, it is obvious that Miller's account can be called *teleological*.

We finish this brief survey with Tuomela (2007) 's account, or rather accounts, since he does not provide one but a whole array of definitions of various joint actions. Tuomela does not oppose the previous approaches but integrates and complements them. He also brings in several original features, among which the main one is what he dubs the we-mode. Even if there are different levels at which an action can be called social, all explored by Tuomela, he holds that to be fully joint, an action has to be accomplished in the we-mode. What is it exactly? "Acting as a group member in the we-mode sense *constitutively* involves acting for a *collectively constructed group reason* – the group gives a group member reasons to think, "emote" and act in certain ways." (Tuomela, 2007, p.3.). Like Miller, Tuomela puts group reasons – in other words, collective goals – at the center of joint action. He provides a definition of joint action that shares the structure of the previously described ones: a set of individual intentions to act to realise a collective goal, which are common knowledge, and the fact that they are common knowledge partly justify them. So there is no real surprise here. But in his view, collective goals are not necessarily defined in terms of individual ones. Rather, they can stem from a collective process - they are discussed and agreed upon by group members. For lack of a clearer term, Tuomela's account can thus be considered as *constructivist*.

2.2. Weak joint action

For all their differences, accounts of joint action share a common structure. All joint intentions are composed of a set of individual intentions that are justified partially by each other, as well as by other reasons, this whole set of conditions (the epistemic *content*) being common knowledge among agents (the epistemic *structure*). The variety of existing accounts originates in that of the multiple ways by which intentions

can arise. Depending on authors, the focus can be put on coherence constraints, on commitments, on goals and/or on social construction.

Now imagine the following situation. Mary, a French woman, has spent some time in country A and appreciates its culture. She naturally happens to be very concerned by the ongoing conflict between A and another country B. One day, by watching the TV news she learns that a demonstration is to take place a few days later in order to protest against the conflict and urge the French government to step in and take some diplomatic intervention in order to stop it. Mary only knows when and where the demonstration is supposed to start (quite far from where she leaves, actually), but not how many people will participate. On the following day, a man in the street gives her a pamphlet advertising the same event, without containing any more details. Mary decides to make the effort to go. On the demonstration day, she turns up, as well as tens of thousands of others. The demonstration is deemed a success, and the government is led to reconsider the unresponsive stance it has kept so far (happy ending).

When all agents implied are in a situation similar to Mary's, I call this collective action a *weak joint action*. Intuitively, it can be characterized by the fact that the agents' beliefs about others do not come from a unique, common source. Rather, the agents' information depends on a distributed cluster of events or facts. Mary received a pamphlet, others may have received one too, but may instead have seen a poster, heard a radio announcement, watched a TV advertisement, etc. I call this property the *multiplicity of epistemic sources*, which can also be understood as an absence of direct epistemic link between agents: an agent's belief about another one are not necessarily based on a unique event or fact that they both witnessed.

Is this not too weak to even be called a joint action anymore? It seems not. Demonstrations where a significant number of individuals are in such a position are far from being uncommon and can have considerable consequences. Advertising such an event is a reliable enough way to make it happen, even if the success can never be guaranteed in advance. It could be argued that in most demonstrations, things happen quite differently than in my toy example. They are always activists, militants, organisers who have previous knowledge about the context, included the expected number of the participants. Some participants will communicate before the event to ascertain whether others are coming or not; the event may be advertised on an Internet page, where the number of comments could allow anyone to form a rough prediction of its success or failure. I do not claim that people who participate in demonstrations never have any idea whatsoever about the future number of participants, and that this idea cannot influence

their decision to come. However, in some cases, a substantial number of persons still decide to turn up on the basis of scarce and scattered information, without previously contacting any of the other potential participants.[8]

Moreover, there are cases of spontaneous mobilization, as for example in flash mobs. These are very extreme kinds of collective action, where people who do not know each other gather in a predetermined place, perform some predetermined identical action so as to produce a salient result, and then break up. The objective of a flash mob is usually no more than to produce a conspicuous, out of the ordinary and temporary mass effect. Even if the places and the action to be performed are decided beforehand, they are merely advertised, without any additional information as to who may participate or how many people can be expected. That the aim of a flash mob is less important or meaningful than that of most demonstrations does not make it less of a joint action. According to the intuitive definition of weak joint actions, some of their main characteristics are that they are *one-shot interactions*: for as soon as the interaction is sustained or repeated, individuals would have common events at their disposal, that is, common epistemic sources. Moreover, for a weak joint action to be anonymous (even though it is not a necessary property) can usually ensure that participants have gathered their information from distinct epistemic sources.

Why introduce weak joint actions? Because they do not seem to fit the previous analyses of joint action. In weak joint actions, individuals only have one decision to make: to participate or not. Bratman's coherence condition is thus trivially met: no individual action can hinder another.[9] It should also be clear that agents do not have to feel committed to participate in any way. In principle, they may even reconsider the decision to participate, as no one knows they are supposed to come in the first place. Gilbert's normative dimension is thus absent. For most agents, who do not organise a demonstration or a flash mob, there is no collectively constructed goal, so they do not act according to Tuomela's we-mode. Now, it may be argued that weak joint actions at least contain some kind of collective goal: after all, all participants want the demonstration to be a success, which can only be the case if

[8] The point is not that all demonstrations (or in the next section, all flashmobs) are cases of weak joint action; indeed, as means communication grow more ubiquitous, less and less are. However, some of them used to be, and sometimes still are. There are other examples: think of an agent deciding to turn up at a party while knowing only the host, who hasn't told her who else would be present.

[9] Cases where only a very limited number of agents can participate notwithstanding.

enough participate. But non-participants may also want this: if someone wants the demonstration to be a success and happens to be certain that enough participants will turn up, she may not bother to participate herself. Moreover, some participants may turn up just because of some secondary incentive (e.g. eating free food for one day, meeting people, etc.) rather than because they care for the demonstration itself. Still, these answers may sound unconvincing and weak joint actions may not seem to be clear counterexamples to Miller's account yet.[10] Hopefully, the next paragraph's objections will cull this reluctance. The existence of weak joint actions at least casts doubt on the generality of the classical analyses and on the question of whether the conceptual core of joint action has not been misunderstood.

Consequently, weak joint actions escape the existing definitions. However, they may still share the very common structure which classical definitions of joint action enjoy. Recall that the authors we evoked locate the core of jointness in the reasons or causes of the individual intentions, that is, the epistemic *content* of joint actions. But there is another more general source of jointness: the epistemic *structure*, namely the common knowledge condition, which stipulates that all conditions should be public among agents. In particular, for anything to be public within a group, all individuals have to know the list of its members, or at least to be able to characterize it. This can be done in the case of a public event, but not of multiple events. If a public speech is given, the members can be defined as 'speech hearers'. But in weak joint action, such a public event is absent; instead, there are various, possibly dissimilar local events (speeches, pamphlets, posters, etc.). As a consequence, neither the set of participants nor their number is known beforehand.[11] It thus seems that there can be no common knowledge between them, and in particular not about their respective intentions to participate. If this is true, neither the epistemic structure nor the epistemic content of the definitions of joint action fit weak joint action. However, the next two sections suggest that at least the epistemic structure can be preserved, through an amended concept of common knowledge.

[10] On the role of goals, see section 4.1.

[11] It is neither known *de re*, nor *de dicto*. In a public speech, one can form beliefs about others, defined as speech hearers, even if she does not know or perceive them individually. But in the case of multiple events, the problem is that the set of relevant events is unknown.

3. Wither common knowledge?

No definition of joint action can spare the reference to common knowledge. However, as we have seen, common knowledge is apparently absent from weak joint action. Must we do without the concept of common knowledge altogether? It may be unfortunate, since this is only potential common point between weak and classical joint action. The fact that all existing accounts are underpinned by the ubiquitous common knowledge condition bears witness to its importance. The content of knowledge and beliefs already differs between weak and classical joint action; if their structure was different as well, there would be no remaining conceptual similarity between them. If we defined a weak form of joint action that hardly resembles existing definitions. there would be no real guarantee that such an exotic account still concerns joint action, rather than, for example, some mere juxtaposition of individual actions that have salient collective effects. In other words, because there is a degree of arbitrariness in definitions of joint actions, it is important to ensure some *continuity* between them.[12]

3.1. Common knowledge

Let us now turn to the analysis of common knowledge.[13] As we have seen, one problem is that to fit a definition of weak joint action, common knowledge has to be weakened, or simply suppressed. We favour the former option; but to prefer the latter is a more popular attitude, which must be rebuked first.

What is common knowledge? Earlier, we said that to be common knowledge and to be public are equivalent. More precisely, common knowledge is usually defined in the following way. An event or proposition E is common knowledge among agents A and B if and only if A

[12] However, it would be wrong to think that the whole topic of joint action is entirely independent from empirical matters. First, any definition should be *cognitively realistic*: it should only refer to mental properties, attitudes and abilities that human beings actually have. Second, and maybe more interestingly, the search for definitions of joint action stems from their intriguing efficiency. Human beings regularly manage to cooperate in many contexts. Defining joint action consists in pinning down the exact conditions that reliably lead to collective success. Even if the joint action theorist may be tempted to adopt a purely normative stance, by determining a priori when a collective action *should be* called a joint action, the very appeal of the topic lies in the impact joint actions have in reality. Definitions should identify the real mechanisms for joint action, because they are what explain its efficiency – section 4.1 addresses this issue.

[13] Authors generally prefer to speak of common belief, which is slightly different. However, the difference between common belief and common knowledge is irrelevant to the following discussion.

and B both know that E, both know that they both know E, both know that they both know that they both know E, and so on. This can be called an *iterative* definition, and has given rise to one major objection: because common knowledge is equivalent to an infinity of statements, it is cognitively unrealistic. An agent can only be in a situation of common knowledge if he has knows an infinite number of propositions, a task for which her memory and cognitive abilities are insufficient, since they are bounded. Common knowledge could only ever be attainable by artificial constructs that reason faultlessly, are logically omniscient (they draw all possible consequences from the available premises) and have infinite storage capacities. To this objection can be added the first one we mentioned, concerning the role of common knowledge in joint action: it presupposes that agents have knowledge about the knowledge of other individuals, which is impossible in some cases. In other words, the concept of common knowledge is questionable because of two distinct properties: it is idealized, and it based on nested individual knowledge operators. Those who consider common knowledge as idealized argue that is should be abandoned and replaced by a cognitively less demanding notion.

One answer to the idealization objection is the following: common knowledge can receive a non-iterative definition such that it does not suffer from the idealization problem anymore. The idea is to notice that common knowledge often appears when public events happen.[14] For instance, if someone argues in a public speech that death penalty is wrong, it will create common knowledge among the listeners that she thinks death penalty is wrong. Common knowledge of a proposition E can thus be defined by the existence of a public event F, that is, such that whenever F is the case, everyone knows that it is the case, where E is implied by F. Take an event F: if it is public, then everything that is logically implied by F can be said to be common knowledge among agents. The crucial result is then that this definition can be proven to be equivalent to the iterative one,[15] so that there is no conceptual loss. But because it does not imply an infinite number of knowledge statements, it cannot be considered as idealized anymore.

[14] This was actually the line followed by Lewis (1969), who famously introduced the concept. Aumann's (1976) seminal account formalized a similar intuition, as Monderer and Samet (1990) made clear. Despite these early efforts, and although it was criticized, for instance by Clark and Marshall (1981), the iterative definition ended up taking over as the intuitively preferable definition of common knowledge. See Cubitt and Sugden (2003) for a formal reconstruction of Lewis' account, and Paternotte (forthcoming) for a recent effort to do justice to the public event account.

[15] However, the equivalence does not hold in all formal frameworks; see Barwise (1988).

A more satisfactory account of common knowledge is thus the following. Common knowledge does not necessitate an infinite number of knowledge statements, but only implies a finite number of those, which stem from a public event. The remaining knowledge statements are merely potential: they could be known but do not need to be. Common knowledge really appears in a situation that has certain properties (namely, containing a public event) and provides not knowledge about what others know but only good reasons to believe it.[16] As a consequence, common knowledge is often obtained by default: agents may think that something was common knowledge among them (they had compelling reasons to believe it) before realizing that, for instance, one of them had not been, say, properly listening to the public speech. This means that common knowledge is not logically deduced from a set of propositions, but inductively inferred from a situation.

3.2. Criticisms and alternatives

Fundamentally, there are several problems with common knowledge. First, some theorists argue that the classical concept is cognitively unrealistic and/or useless when explaining human collective actions. It should thus be replaced by alternative versions, weakened or even abandoned, even in the usual definitions of joint action. Secondly, is there classical common knowledge in weak joint actions? We already gave a negative answer to this question in section 2.2. Thirdly, it is dubious whether even a weakened version of common knowledge can fit weak joint actions, that is, cases of multiple epistemic sources. After last section's abstract analysis, this section claims that classical common knowledge is not unrealistic, by commenting the views of several theorists who have argued to the contrary, but that it should and could be weakened to fit weak joint actions.

3.2.1. *Dancers and robots*

In a recent text, Parikh (2005) makes the following analysis:

> Suppose than Ann and Bob are dancing and at the moment when Bob puts his right foot forward, Ann must pull her left foot backward. If one action happens without the other then either Bob will fall, or he will step on Anns foot, both of which are undesirable occurrences. But does Bob ask himself, "Does Ann know that I know that she knows that I know that she is going to pull her left foot backward?" Of course not. What does happen is that they are both listening to the same music and if they have already trained themselves to dance properly to this music, then their movements will

[16] Again, this follow Lewis' original analysis (1969, p.56), formulated in terms of 'reason to believe' and 'indication'.

> be in concord, and co-ordinated action will take place without conscious common knowledge (and is there any other kind?). If they are not trained to dance the particular dance, then no amount of common knowledge will help them.

This text gathers many problems faced when interpreting common knowledge. It considers common knowledge on the basis of the iterative definition, stating that an infinite number of statements should be known explicitly. The previous section suggested that common knowledge can as well be defined by a public event that involves the knowledge of only a few things. Consequently, common knowledge exists and is not entirely 'conscious': the infinite number of epistemic iterations could be deduced by an agent with unbounded cognitive power, but common knowledge can exist without these deductions being made. In this sense, common knowledge generally exists by default: agents only start checking what they know about each other when they doubt that there is common knowledge among them. If human dancers stumble or make a wrong move, they will start wondering whether the sequence of steps is really common knowledge between them. But as long as they are coordinated, common knowledge holds: more precisely, as long as the dance goes on smoothly, it provides no reason to think that common knowledge does not hold. This common knowledge comes from the fact that dancers have received a similar training, have learnt the same sequence of steps and know this.[17]

Even if dancers had no training, asserting that "no amount of common knowledge will help them" is ambiguous at best. They could have obtained common knowledge when studying a book together without trying to practice; if they are gifted, they should be able to dance (but this is a problem of accomplishing complex actions which has nothing to do with coordination of multiple agents); or common knowledge could arise during the interaction, similarly to Lewis' example of two rowers learning to coordinate on the fly. Two dancers not knowing each other may still discover during their first dance that their styles fit perfectly; from an initial trial period, common knowledge would then appear. But if the sentence "no amount of common knowledge will help them" refers to epistemic iterations which would be provided successively, the statement is partly true: dancers will receive no help - but then what they are receiving is not common knowledge.

Later in the same text, Parikh claims that nothing would change if Bob was actually dancing with a robot, a case where no common knowledge exists. This only proves that common knowledge is not

[17] In Lewis' vocabulary, there is a precedent that makes salient a way to coordinate.

necessary for successful coordination - and that its existence cannot be deduced from the mere observation of behaviors. From the dancers' point of view, there are no proofs of common knowledge but only hints depending on the observed events. Moreover, this case does not even qualify as a joint action at all.

Overall, Parikh intends to cast doubt on the idea that agents use common knowledge even in cases of strong joint action. However, what makes it unlikely is that he uses a mistaken interpretation of common knowledge by assuming that it implies an infinite number of knowledge statements, of which agents should be simultaneously aware.

3.2.2. *Children and animals*

We now turn to the most common criticism regarding common knowledge. According to Tollefsen (2005), the "common knowledge problem" - that is, that common knowledge requires too strong cognitive abilities - is one of the two major problems faced by attempts to define joint intention. Similarly, Peacocke (2005) recently claimed that: "a range of phenomena that have been characterized in terms of mutual knowledge should rather be elucidated in terms of joint attention", where joint attention between agents is defined as open perceptual knowledge, itself basically defined as follows: agents both perceive a fact, are aware of this shared perception and aware of this general state of awareness. This account is quite similar to Gilbert's account, the reference to ideal agents notwithstanding. Tollefsen herself agrees that the use of this notion would for example fit young children, who are able to coordinate though unable to perceive the beliefs or intentions of others.

This position is not mistaken in itself, but replacing common knowledge by another concept can easily be done for the wrong reasons and thus actually prove useless. The justification is usually made in three steps: (1) Common knowledge is condemned for being cognitively unrealistic; (2) less demanding but conceptually equivalent concept is built; (3) the latter can replace the former in an analytical definition of a social notion.

As I already argued, (1) is false when it is based on an iterative definition of common knowledge - which is Peacocke's case (2005, p. 300). There is nothing wrong in stating that ordinary human adults are in a state of common knowledge. Nonetheless, (1) can be justified in cases where the agents do not even have an elaborate enough theory of mind to, for example, form beliefs about each other's beliefs (which is needed even in a public event account of common knowledge), as in the case of young children or even animals. This is an argument of Tollefsen (2005). The same replacement concept can thus be used

for the wrong reasons as well as for the right one - the two following remarks do not apply to this latter case.

As for (2), most often common knowledge and the concept built to replace it are not equivalent on the conceptual level. Consequently, such a strategy actually confuses two different aspects of common knowledge. Barwise (1988) famously distinguished three distinct problems concerning common knowledge: one can try to define it, to find out in which conditions it appears, or to explore the way in which agents use it. Peacocke's move consists in replacing common knowledge by sufficient conditions for common knowledge to appear, that is, in tackling the second problem. Common perception or joint attention, mutual recognition, eye meeting... are all concrete phenomena that lead ordinary agents to common knowledge. In fact, the replacement concepts are generally causes of common knowledge. What happens in (3) then, is that common knowledge is replaced by something entailing it - that is to say, common knowledge is not really replaced and there is no reason to think it should be abandoned. In particular, the resulting improved definition of joint intention requires nothing less from agents than before the replacement. Consequently, the replacement does no harm but cuts no ice. In other words, in general the new definition will only seem more realistic because common knowledge falsely seemed unrealistic.

3.2.3. *Manifestness and openness*

Most authors interested in joint action have not really discussed common knowledge and endorsed its usual, iterative definition.[18] Somewhat ironically, although the previous criticisms of common knowledge are fairly recent, a few weakened versions of it had already been advocated by philosophers more than twenty years ago. We now give two such contributions from philosophers of social science, which will illustrate the idea that common knowledge could be slightly modified rather than eliminated.

Sperber and Wilson (1986) were early to suspect that common knowledge could be an idealized concept, and to try to replace it by a more realistic notion. They introduced a concept of *mutual manifestness*, defined as follows. In a given situation, an agent's cognitive environment refers to the set of facts which are manifest for her: all she may perceive or come to know - whether she perceives it or not is a contingent matter. For a group of agents, a shared cognitive environment is a cognitive environment where the identities of the agents are manifest. At last, a fact is said to be mutually manifest if it belongs to

[18] Cf. for example Bratman (1992), p. 335, note 15; Miller (2001), p. 59.

a group's shared cognitive environment. This definition is quite close to the account we sketched earlier: it says nothing about agents' actual beliefs and basically describes a situation rather than mental states. But a problem arises because for a fact to be manifest, it is sufficient for it to be perceptible. For example, it would be mutually manifest among a group that there is a elephant-shaped cloud in the sky, even if no one actually noticed it. As a result, mutual manifestness is to weak a notion to replace common knowledge.

A more successful attempt is that of Gilbert (1989), who provides a definition of weakened common knowledge. Gilbert describes a "paradigm situation of common knowledge" of a statement E, as the conjunction of the following facts: Agents are normal human beings with normal perceptual organs, normal reasoning capacities and similar conceptual equipment; they perceive that; they perceive each other; E holds; agents perceive that E; they perceive each other perceiving E; E is open*; agents have the concept of openness* and know they all have it; and they perceive that E is open*. Gilbert defines *openness** by introducing the notion of a smooth reasoner counterpart, which is like a human being "but whose reasoning is untrammeled by limits of time, memory capacity, and perseverance", that is, able to infer every logical consequence of any situation. E is then said to be open* to agents when their smooth reasoner counterparts would be able to reach classical common knowledge of E.[19]

Such an approach is satisfying because it does not fall prey to the temptation of the iterative definition. Gilbert's account is based on a public event, and as a result is quite similar to the description I have given in the end of section 3.2. The very same situation leads normal agents to ordinary common knowledge and ideal agents to iterative common knowledge. In other words, what she calls openness* is actually what I have been calling (realistic) common knowledge.[20] This shows that there is no need to eliminate common knowledge from definitions of joint action. All that remains is a terminological dispute.

3.3. Indirect common knowledge

Imagine that Mary, a philosopher, takes the train for a long trip. At one point, out of boredom, she starts talking to the person next to her, who she has never met before, and who when asked about his job reveals he is a philosopher too. It gives Mary a vast amount of knowledge about him and possible topics of conversation. For example, she knows she

[19] Note that this account is very close to Lewis'.

[20] Gilbert's account is still unsatisfying for other reasons that do not matter here (Paternotte, forthcoming).

could probably start talking about Kant or Aristotle without having to explain who they are first - and the man knows that too. Realizing they are both philosophers makes them reveal their common background, in other words creates common knowledge between them. But they have never studied together, so no public event explains this. What they have instead in his case is what can be called *indirect* common knowledge.

Why is indirect common knowledge still common knowledge? Because it can be defined in a very similar way. Direct common knowledge is based on a *public* event, and indirect common knowledge on *typical* events. A public event gives agents who perceive it knowledge about what the others perceive. Typical events are sets of similar events, each of which gives agents knowledge about those who perceive any of them.[21] There is common knowledge among mathematicians about what a differential equation is, not because they all took the same course when they were students, but because several course gave them information about what *any student having received a similar education* would know. In a way, a public event is scattered and replaced by a set of disjoints typical events. The more one hears about Kant during her philosophy studies, the more likely she thinks it is that any other philosopher would have encountered circumstances in which he would have heard the name as well.[22] In particular, examples of typical events include very common phenomena like, say, television advertisements – when you watch one advertisement spot, you know that they have been and will be others, and that people watching them will get the same information you just got. Cubitt and Sugden (2003) precisely defined indirect common knowledge[23] along those lines and showed that it is common knowledge indeed, because it can lead to the same amount of knowledge; only its source differs.

As a result, in weak joint action, provided they have *typical enough* experiences - experiences that make one another likely - agents who have no *common* experience can still have indirect common knowledge. But common knowledge of what exactly? In other words, even if we

[21] More precisely, a public event is defined as such that if an agent knows it, she knows that others know it. Accordingly, a set of typical events is such each event of the set leads to some identical knowledge, and that whenever each agent knows one event, she will know what others know another one. See Cubitt and Sugden (2003), p.191.

[22] Indirect common knowledge is closely related to Stalnaker (2002)'s concept of common ground, corresponding to 'presumed background information' (p.701), that is, to a belief that something is common belief. However, common ground does not have to be true, but indirect common knowledge has. Moreover, Stalnaker focuses on how something can become common ground, through what he calls accommodation, and only static indirect common knowledge interests us here.

[23] In their vocabulary, it would be 'distributed' rather than indirect (p.192).

know *how* mental attitudes are shared between agents, we have to determine *what* is shared - what is the content of these attitudes. This is the topic of the next and last section.

4. On the psychological side

We have seen that even in weak joint action, the common knowledge condition can be salvaged. One source of jointness between agents is still present. But this does not suffice to identify what is really shared in weak joint action: common knowledge may be a necessary condition, but it is far from being sufficient. After all, groups of people can have common knowledge of various things without accomplishing any joint action. Individual action depends of an agent's beliefs (or knowledge), but one can hold various beliefs without necessarily acting. Mutatis mutandis, the same goes for joint action and common belief (or knowledge). Consequently, we now briefly turn to psychological explanations of joint action.

4.1. Propensities to adopt goals

What kind of beliefs can lead people to cooperate in situations of anonymous, one-shot interactions? Several surveys (Dawes 1980, Orbell and Dawes 1981, Kollock 1998, Tuomela 2000, chap. 12.) concerning the factors that influence cooperative behaviour reveal that not many psychological mechanisms are likely to explain them. Kollock's survey mentions only two: social value orientation (p.192) and group identity (p.194). More recently, Bicchieri noticed that "only group identity and social norms have not been eliminated by experimentation as possible explanations" (Bicchieri, 2006, p. 145.).[24] Basically, explanations based on social value orientation claim that individuals facing a cooperative dilemma may adopt 'social' preferences, which are distinct from the situation's material payoffs. According to group identity explanations, when the similarity of individuals is made salient, they may start perceiving themselves as members of a group, and as a result adopt its preferences, reason collectively, favour ingroup members over outgroup ones, etc.[25] Social norm accounts argue that situations offering a po-

[24] Bicchieri asserts this when discussing the effects of communication on cooperation. However, her account of social norms can be applied to situations with no communication as well.

[25] Numerous studies (For instance Tajfel et al. 1970, Tajfel 1973, Dawes et al. 1988, Brewer and Gardner 1996) have shown that many apparently innocuous factors can trigger group identification – from arbitrarily dividing agents into groups to

tential mutual benefit trigger a preference change in norm-sensitive individuals (Bicchieri 2006, p.3).

This is not the place to adjudicate between these mechanisms. However, for our purposes it is sufficient to note that each explanation postulates a transformation of individual preferences that brings them closer to some group interest.[26] Such transformations depend both on salient characteristics of the situation that trigger them, and on the individuals' sensitivity to them. In other words, even if an individual is sensitive to the relevant cues (e.g. the possibility of a collective benefit), she cannot be certain that it will affect others; she can only reason from others' estimated tendency, or propensity, to adopt such 'social' preferences too. As a result, in cases of weak joint action, what can be common knowledge will be the individuals' *propensity to adopt social preferences* (or collective goals), determined from objective features of the situation.[27] Competing explanations only disagree as to what context triggers what transformation; but not on the fact that preference transformations are necessary. Overall, bringing in tendencies or propensities departs from classical joint action, in which the content of knowledge and beliefs concerns individual mental attitudes.

Typical events such as sets of advertisements, announcements, speeches, etc., that concern a collective outcome are what can create indirect common knowledge about the fact that people are likely to adopt the outcome as a goal.[28] This in turn depends on several factors: the estimated impact of the announcements and advertisements, the more or less convincing way in which they deliver the message, the attractiveness of a collective outcome, the past joint actions (whether they succeeded or failed) that agents have witnessed or of which they have been part, etc. That all this information plays a role does not necessarily mean that agents perform complex calculations, but only that every piece of it can influence the decision to participate or not.

---

having them read texts and circle the collective pronouns ("we", "us", etc.) – and its relevance to cooperation has been recently emphasized (Bacharach, 2006).

[26] In what follows, I only consider social norms and group identity, as they both entail a preference transformation, that is, a social value orientation.

[27] In cooperative contexts, I consider as equivalent the adoption of social preferences or of a collective goal. This has no influence on the later discussion, and is justified by the following facts: in social norms explanation, only a situation containing a collective outcome (a mutual benefit) can trigger the change of preferences (Bicchieri (2006, p.3); as for group identification, recent empirical studies (Cremer and Van Vugt 1999, de Cremer et al. 2008) have shown that it is better explained as a change of goals.

[28] Note that this does not make weak joint action equivalent to Tuomela's we-mode, which, as we have seen, requires that collective interest, or collective goals, be collectively built. Here on the contrary, the goals can directly emerge from a situation.

People can easily adopt preferences or goals, and to a certain extent expect others to do the same as well: this is how indirect common knowledge can appear and then sustain weak joint action.[29]

4.2. Of ants and men

I now turn to some more general considerations related to a general account of weak joint action. First, let us summarize it. In a weak joint action, agents do not know each other and interact only once. They act partly because they adopt a salient collective outcome. The main link between agents stems from the degree of commonality that they estimate their experience and intuitions have. An agent may not know what others think or know, but she has at least some vague idea as to how public or well-known her knowledge is. She can guess the average profile of those likely to think similarly to her, or to have the same information and maybe the same interests. She can also get an impression of how many people are likely to be influenced by a public message advertising the potential joint action, depending on how widely spread the message appears to be. She can thus form a rough estimate as to how many people could join. At last, depending on this estimation and on whether she identifies with the potential group of participants, she may or may not participate. Surely, this is all guesswork, but what else could it be when no information about others is available?[30]

Two clarifying remarks are now in order. First, the combination in the same definition of indirect common knowledge and of propensities to adopt a goal is not artificial. As a matter of fact, these propensities could not hardly play a role in weak joint actions if common knowledge was absent. To adopt a group's interests does not automatically entail the participation to any joint action. Only when an agent believes that enough others will reason similarly to her and participate, or that

[29] A complete analysis should include a game-theoretic framework, in order to compare weak joint action to Nash equilibria (for instance) and to work out the links between a collective goal and individual preferences. Addressing these issues here would have us wander too far from this paper's focus, namely the epistemic dimension of weak joint action. For a game-theoretic framework based on group identification, see Bacharach (1999).

[30] All this seems to mean that common knowledge comes in degrees. In my opinion, which is consistent with Lewis' (1969), common knowledge is an extreme kind of common belief, but the difference between the two is hard to pinpoint precisely - it depends on what maximum degree of belief about others' belief humans can reach. Moreover, in the joint action literature, theorists usually only talk of common belief without being too clear about its degree. A definition of weak joint action, like other joint actions, should accommodate cases of partial common belief as well. This does not impinge on the rest of the analysis: whatever its degree, common belief shares common knowledge's structure and content.

others in general are likely enough to participate, will she decide to participate herself. In other words, propensities of goal adoption *have to* be supplemented by (indirect) common knowledge in order to trigger a decision to participate. And this point helps see more clearly how weak joint action really is a kind of joint action, which was an initial worry of ours. Think of a case à la Tuomela, based on a collectively built goal and some common belief. Similarly, weak joint action is based on what we could call weak collective goals (salient collective interests) and indirect common knowledge. In other words, weak joint action is obtained from joint action by adequately weakening its two core conceptual parts.

Second, it seems that weak joint action lies in the interval between full joint action and simpler cases of purely behavioural collective actions such as found in animal species.[31] Think of collective actions in an ant colony: amazing coordination levels and collective outcomes are reached whereas the ants do not reason about each other's mental states. Somewhat similarly in weak joint action, agents do not reason from specific mental states, but anticipate tendencies for the average individual to act, or numbers of potential participants; they consider the set of other agents as an independent whole, the behaviour of which they tentatively predict. For this reason, it is tempting to see weak joint action as minimal joint action, that is, as the simplest form of joint action that can be accomplished by humans, or more precisely as the exact bound that marks where purely behavioural collective action ends and human cooperation begins. According to this view, if we go from ants to men on the scale of continuously increasing cognitive abilities, the amount of what can be shared between individuals would linearly increase as well, and we would also go linearly from the simplest collective actions to the most complex ones. Reaching weak collective actions would then mark the beginning of the realm of human joint action.

However, I think this is view is mistaken. This general tendency to consider that the more sophisticated the species, the more complex the features that constitute joint action, is not necessarily true. Joint action *can* grow increasingly intricate as cognitive capacity gain complexity, but does not have to. Indeed, human beings, whose strong cognitive abilities allow them to sophisticated kinds of joints action, can also occasionally stoop back to lower, less refined forms. But weak joint action is not one of those: it is a form of disconnected joint action that is minimal in terms of what is shared *by way of common perception and mutual interaction*. This does not mean that what is shared would

[31] By purely behavioural collective action, I mean any collective action that has some observable results, regardless of the jointness that may exist between individuals.

be accessible for individuals with lesser cognitive capacities. Indeed, sharing expectations that admit degrees, making inferences about other people's expectations while having no further information about them is far from being a straightforward task. In other words, only with high enough cognitive power is it possible to still act jointly on the basis of a limited amount of indirect evidence about others. Human beings can still coordinate in very abstract ways, which only seemingly mirror the forms of automatic coordination found in animals such as ants. The idea that the complexity of joint actions lies in the intricacy of nested mental attitudes is misguided; it also lies in the nature of their content. That joint action is still possible when no common event or fact unites individuals is proof of their high cognitive abilities.

5. Conclusion

I have characterized weak joint action as a kind of minimal joint action, in which individuals' knowledge and beliefs come from multiple epistemic sources, for instance some cases of one-shot interactions, and have analyzed their epistemic dimension. Although the structure of joint action definitions usually consists of individual mental attitudes that are common knowledge, I have argued that these aspects are absent from weak joint action. Nonetheless, I have shown that there is still indirect common knowledge - based on the typicity of events rather than on their publicity - and that its content concerns tendencies or propensities to adopt certain group preferences or certain collective goals. As a result, much of the classical structure of joint actions, although with weakened components, is preserved in weak joint action. However, the increased complexity of its epistemic contents make it cognitively demanding. Humans (at least) are capable of joint action even in bleak situations, in which the sophistication of their epistemic content can make up for the group's impoverished epistemic structure.

References

Aumann, R.J. Agreeing to Disagree. *The Annals of Statistics*, 4(6):1236–1239, 1976.

Bacharach, M. Interactive team reasoning : A contribution to the theory of cooperation. *Research in Economics*, 53:117–147, 1999.

Bacharach, M. *Beyond Individual Choice : Teams and Frames in Game Theory*. N. Gold and R. Sugden, editors. Princeton University Press, 2006.

Barwise, J. Three Views of Common Knowledge. In M.Y. Vardi, editor, *Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning about Knowledge*, 365–379, 1988. Morgan Kaufmann Publihers Inc., San Francisco.

Bicchieri, C. *The Grammar of Society - The Nature and Dynamics of Social Norms*. Cambridge University Press, Cambridge. 2006.

Bratman, M. Shared cooperate activity. *The Philosophical Review*, 101(2):327–341, 1992.

Bratman, M. Shared Intention. *Ethics*, 104:97–113, 1993.

Brewer, M.B. and Gardner, W. Who is this "we" ? Levels of collective identity and self identification. *Journal of Personality and Social Psychology*, 71:83–93, 1996.

Camerer, C. *Behavioral Game Theory*. Princeton University Press, 2003.

Clark, H. and C. Marshall, C. Definite Reference and Mutual Knowledge. In A. Joshi and B. Webber and I. Sag, editors, *Elements of Discourse Understanding*, 10–63, 1981. Cambridge University Press.

Cubitt, R. and Sugden, R. Common Knowledge, salience and convention : a reconstruction of David Lewis's game theory. *Economics and Philosophy*, 19:175–210, 2003.

Dawes, R.M. Social Dilemmas. *Annual Review of Psychology*, 31:169–193, 1980.

Dawes, R.M and van de Kragt, A.J.C and Orbell, J.M. Not me or thee but we : the importance of group identity in eliciting cooperation in dilemma situations : Experimental manipulations. *Acta Psychologica*, 68:83–97, 1988.

De Cremer, D. and van Knippenberg, D. and, van Dijk, E. and van Leeuwen, E. Cooperating if one's Goal are Collective-Based - Social Identification Effects in Social Dilemmas as a Function of Goal Transformations. *Journal of Applied Social Psychology*, 38(6):1562–1579, 2008.

De Cremer, D. and Van Vugt, M. Social identification effects in social dilemmas : A transformation of motives. *European Journal of Social Psychology*, 29:871–893, 1999.

Gilbert, M. *On social facts*. Princeton University Press, 1989.

Kollock, P. Social Dilemmas : The Anatomy of Cooperation. *Annual Review of Sociology*, 24:183–214, 1998.

Lewis, D. *Convention - A Philosophical Study*. Harvard University Press, 1969.

Miller, S. *Social Action - A Teleogical Account*. Cambridge University Press, 2001.

Monderer, D. and Samet, D. Approximating common knowledge with common beliefs. *Games and Economic Behavior*, 1:170–190, 1989.

Orbell, J.M. and Dawes, R. Social Dilemmas. In G.M. Stephenson and J.M. Davis*Progress in Applied Social Psychology*, vol. I, 37–65, 1981.

Pacherie, E. Is collective intentionality really primitive ?. In M. Beaney and C. Penco and M. Vignolo, editors, *Mental processes: representing and inferring, Cambridge: Cambridge Scholars Press*, 153–175.

Parikh, R. Logical Omniscience and Common Knowledge : WHAT do we know and what do WE know ?. In R. van der Meyden, editor, *Proceedings of the 10nth Conference on Theoretical Aspects of Rationality and Knowledge*, 62–77, 2005. National University of Singapore

Paternotte, C. Being Realistic about Common Knowledge: a Lewisian Approach. *Synthese*, DOI:10.1007/s11229-010-9770-y, forthcoming.

Peacocke, C. Joint Attention : Its Nature, Reflexivity and Relation to Common Knowledge. In N. Eilan and C. Hoerl and T. McCormack and J. Roessler, editors, *Joint Attention : Communication and Other Minds*, 2005. Clarendon Press, Oxford.

Searle, J. Collective Intentions and Actions. In P. Cohen and J. Morgan and M.E. Pollack, editors, *Intentions in Communication*, M.I.T. Press, 1990 .

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

Sperber, D. and Wilson, D. *Relevance: Communication and Cognition*. Blackwell Publishing, 1986.

Stalnaker, R. Common Ground. *Linguistics and Philosophy*, 25:701–721, 2002. Linguistics and Philosophy 25: 701721, 2002.

Tajfel, H. The Roots of Prejudice : Cognitive Aspects. In P. Watson, editor, *Psychology and Race*, Aldine Publishing Company, 1973.

Tajfel, H. and Billig, M. and Bundy, R. and Flament, C. Social Categorization in Group Behavior. *European Journal of Social Psychology*, 1:149–178, 1971.

Tollefsen, D. Let's Pretend! Children and joint action. *Philosophy of the social sciences*, 35(1):75–97, 2005.

Tuomela, R. *Cooperation - A philosophical study*. Kluwer, 2000.

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