p. 2 “To illustrate what is at stake...“

I wasn’t fully convinced that referring to individual intentions helps to make a strong case for considering joint action without shared intention. You argue that individual action can occur without intention and that “understanding intentional action requires understanding at least some non-intentional kinds of action”. The analogy to joint action is that joint action can occur without shared intention. However, I think this analogy may be problematic because the actions you get without intention in the individual case are qualitatively different from the actions you get when individuals act with intentions (i.e. reflexive behaviour is different from intentional action, e.g., jumping for joy is likely different from jumping to exercise in terms of kinematics etc.). In contrast, your case studies (e.g., Ayesha and Beatrice) suggest that joint actions can unfold in the same way with and without shared intentions, and the same outcomes can result from joint action with and without shared intentions. [hedged]

p. 5 case study - suprise

Is it really crucial for this example that Stan and Mavis did not intend to kiss each other? It seems to me that you could make the point about the lack of shared intention even when granting Stan and Mavis the intention to kiss. Isn’t it sufficient for the argument to assume that they are not aware of the other’s intention to kiss? There is still surprise when they kiss because Stan had no idea Mavis wanted to kiss him and vice versa. I think this would make the example more intuitive. [would make the suggested change but deleted the case study]

p. 6 case study - environment

This is a brilliant case study. I think the example would work even better if it were less artificial. A lever than can be pulled down to lift up an object does not seem to have a natural, obvious affordance (one could find out about the consequence of moving the lever by trial and error, but upon seeing the lever for the first time, one would not necessarily know what to do with it). In Melis et al.’s rope pulling study, the rope has a clear affordance because ropes are there to be pulled. I guess replacing the levers with ropes in your example would work to illustrate the contribution of affordances more clearly. And perhaps it would be good to mention that both agents see the consequence of their own/the joint action (some readers might think that the stone wall precludes Ayesha and Beatrice from seeing the heavy block). [all done]

p. 8 “It is now well established…” I would delete “simulating” in this sentence because it sounds as if simulating was a process occurring independently of observation. (You could replace it with “imagining” but I don’t think that’s necessary here). [done]

p. 8/9 The way you brought in motor simulation in Sam and Ahmed’s joint action is very clever. An open question is what exactly Sam and Ahmed simulate/predict: Do they predict the other’s actions, or the consequences of their combined actions? (case study: we are rowing a boat together using one paddle each; do I predict your paddling actions, or the combined outcome of our joint paddling; or both?) It may be beyond the scope of the paper to discuss this, but it is perhaps relevant because even if agents predict joint outcomes, this need not involve shared intentions (see below). [\*\*\*NEEDS MORE THOUGHT. The prediction of a joint outcome doesn’t apply to the Sam/Ahmed case because each needs to predict the others action or its outcome only; so there is perhaps a further issue about predicting the outcome of participatory activities vs. predicting the activities themselves.]

p. 9, last paragraph: Another objection one could put forward is that forming predictive models that include other agents requires shared intentions. I think this can also be countered, e.g., Ayesha and Beatrice could make predictions about the outcome of their combined forces while still treating each other as tools. [done---but reply to this objection occurs earlier, in the discussion of the motor simulation because in the experiments observers produce predictive models although there is no reason to think they have shared intentions as they are not asked to participate]

p. 19 case study-the imaginary we

I was not so convinced by this case study but I find it difficult to say why. I guess the main reason is that unlike in your other case studies, it is hard to imagine finding oneself in this situation. Although logically possible, it somehow lacks “street cred” ☺ I wonder if there is a way to include a more concrete example? If it’s not possible to come up with an example then this might be an indication that this is more of a thought experiment than an actual possibility. As such, it may still do the job but it seems less intriguing than the other cases studies, perhaps because it is more about intention attribution than on-line joint action performance.

[\*\*\*still thinking---more visceral case would involve pantomime horse but harder to maintain innocence about whether think of oneself as engaged in a collective activity; the argument would have to be that although such awareness can hardly be avoided, agents are not acting *because* they have that awareness---which would be hard to argue for.].

The conclusions are great!!

Minor points:

p. 4, last paragraph: Do you really mean “enjoy” or simply “have”? [done]

p. 5: “While this assumption does not hold on every account of shared intention, the variety of cases …” As a non-philosopher I was lost here. This is probably just an issue about wording. [done]

p. 5: “Given that singing and tangoing supply paradigm joint actions, … “ It doesn’t really seem to follow from this that kissing is a joint action. Again, just a wording issue I think. [cut]

p. 8, line 18: gaze patterns were similar … “to” is missing [done]