**Extending the Extended Mind: The Phenomenon of WE-ness**

**Abstract**

Advocates of cognitive extension argue that the human mind super-sizes itself by *embodying* itself in a body, *embedding* itself in an epistemically agential environment and uniting itself with both in *extended* cognitive agency. Call this the 3E-ness thesis. In this paper, I propose a strong version of 3E-ness, WE-ness: In some instances super-sizing results is the creation of a plural subject, a WE. I outline the ontological lineaments of WE-ness distinguishing it from other types of 3E-ness and suggest an evolutionary biological model of its origin based on the emergence of multi-cellular life from single celled-life. And I then turn to some findings in developmental psychology concerning we-intentionality and its features of normative and supra-personal intentionality. Finally, on the basis of these findings, I indicate briefly why a WE-ness account of group agency is superior to two leading competitors, summative and transcendental social constructionist accounts.

(Number of Words in Abstract, 145)

## I. Introduction

Advocates of the extended mind have argued that human cognitive and agential capacities are not only *embodied* and move beyond one's skin into the external environment, *embedding* themselves in the environment, but also are *extended* into that environment so that a thinking, cognizing entity is constituted not only by the embodied mind but also by an embedding environment.[1] Brain, body and environment form a thinking, cognizing entity. The mind has super-sized itself! Consider, for instance, the use of notebooks, computers, the Internet and I phones and situations in which you and friends make plans for the evening. Our world is increasingly a world of embodied, embedded, and extended epistemic agential systems. Call this phenomenon 3E-ness.[2]

In this paper, I propose and argue for a version of 3E-ness that to my knowledge advocates of 3E-ness have paid less attention to, that is, the phenomenon of WE-ness, the extension of cognitive agential agents to form a plural agent, a WE.[3] This is surprising for at least two reasons. Common sense and ordinary discourse abound with talk of plural agents -- corporations and nations, we and us, and they and them.[4] Moreover, there is a highly developed analytic philosophical literature on social action, plural subjects and joint intentionality.[5]

To motivate the discussion of WE-ness, I start with Clark and Chalmers well known thought experiment concerning cognitive extension, presenting an until now unrevealed back-story on Otto and his notebook. Next, I sketch a scientifically based ontology that includes cognitive agency. From there I outline a way to understand WE-ness based on a biologically based model of the origin of multi-cellular life from single-celled life. I then argue that recent findings and theories in developmental psychology concerning we-intentionality and its features

of normative and supra-personal intentionality provide *prima facie* support for the WE-ness hypothesis relative to two of its competitors, summative individualism and transcendental social constructionism.

### II. The Phenomenon of WE-ness: The Back Story on Otto and His Notebook

As Clark and Chalmers tell it, Otto is suffering from an impaired memory. Attempting to remedy this problem, Otto writes down important items in his notebook, for instance, the location of the museum that he wishes to visit. In contrast, Inga just remembers its address. Clark and Chalmers argue that Otto and his notebook form an active epistemic unit. Otto's cognitive activity is extended in the notebook in so far as it plays the same role as Inga's memory and originally Otto's own memory.[6] His is a super-sized mind. But, this is not the only story of super-sizing the mind that we can tell.[7]

In the back-story, we find that the tale of Otto's embodied, embedded and extended mind is richer and sadder than Chalmers and Clark have imagined.[8] Otto and Inga, now revealed to be Otto's wife, moved into a retirement center ten years ago. Four years ago, Otto needed to transfer to the assisted living facility and now he is in the cognitive impairment wing, where Inga visits him everyday and his children pay him frequent visits. Friends, family and Inga have noticed with increasing sadness Otto's deterioration. Otto had been a well-known astronomer, who worked at NASA for many years. He was member of the scientific team that enabled the moonwalk expedition and was a top scientist on an unmanned expedition to Mars. He and Inga had been active in several NGOs and had a wide circle of friends. Otto became unable to do any scientific work some six years or so ago. Their circle of friends has diminished significantly.

Otto no longer recognizes many of his friends.  Indeed, he thinks that the few that he does recognize are always angry with him.

The lesson of the back-story is clear.  Otto's cognitive capacities in their prime were extended well beyond the minimal kind of extension involved in using a notebook as a memory device.  As a scientist and a NGO member he was engaged with others in common cognitively based enterprises that made him a part of a larger cognitive whole whose ends and means of accomplishment were beyond those of its individual parts.  Otto's mind is in a state of painful contraction.  His 3E-ness, especially his WE-ness, is in sharp decline.

### III. The Ontology of WE-ness

Consider a scientifically based ontology in which WE-ness finds its place.[9]  First distinguish between entities with intrinsic ends and entities with merely extrinsic ends.  The latter can serve only as instruments of some entity with intrinsic ends.  I take living things to be primary examples of entities with intrinsic ends.  Such entities have components that serve various functions.[10]  Some living entities have cognitive capacities.  I take a cognitive capacity to be a kind of functional capacity that enables representations of both external and internal reality and that generally results from spatial and temporal environmental complexity.[11]  Only these entities have the capacity to achieve 3E-ness.  Human agents have the capacity to achieve the We-ness form of 3E-ness.[12]  Primary examples of WE-ness are unities of human agents acting together to achieve common ends, ranging from mother and infant to scientific research teams, corporations, governments and non-profit organizations.

We can get a handle on the difference between Clark's individualistic 3Es and WE-ness by considering three different units of selection proposed in evolution theory: gene, organism

4

and group.[13]  Richard Dawkins has proposed that genes are the units of selection.[14]  In addition,

he has maintained that the body that encases the genes be considered as the genes' vehicle.  He

has also argued that extra-bodily environments on occasion function as the genes' extended

phenotype.  Genes, vehicle and extended phenotype form operational unities that achieve various

effects some of which positively affect the genes' fitness and thus their survival and continuance

in further generations, along with in some cases that of the extended phenotype.  Similarly,

Clark's 3E-ness reflects this gene-centered biological model.[15]  The naked brain is embodied in

an embedded body that joins with the embedding environment to produce cognitive effects.

In contrast, biologists have traditionally considered organisms as the units of selection.[16]

Organisms have a variety of traits, some better adapted to the environment than those of other

organisms.  A given environment selects for some organisms over others, yielding them

relatively superior fitness.  Consequently, they survive better and reproduce more than their

competitors.  Organisms are the units of selection because they interact with the environment and

their fit with the environment determines their evolutionary fate.  But organisms also act on their

environments shaping them to their evolutionary benefit.  They too create extended phenotypes.

Thus, on the traditional view, individual organisms embody and embed genes and frequently

become extended.[17]  However, besides genetic and individual selection there is group selection.

Though controversial during a portion of the second half of the twentieth century, evolutionary

theorists now not only acknowledge it as a genuine theoretical possibility but also as a plausible

empirical reality.  The major transitions in life history, for instance, the change from one-celled

organisms to multi-celled organisms, offer promising instantiations of group selection.[18]  Such

group selection involves group-level differences, selection on groups with those differences and

differential survival and reproduction of groups.  I contend that selection on the levels of

organism and group provide a better model for understanding and explaining WE-ness forms of 3E-ness and it's origin than does Clark's implicit gene-centered model.

### IV. The Origin of WE-ness

In a recent paper Samir Okasha has attempted to formulate an evolutionary criterion for the movement from individual level selection to multi-level selection.[19] Okasha illustrates the criterion for the emergence of multi-levels by considering a case of the emergence of a multi-celled organism from a single celled one using the work of evolutionary biologist, Rick Michod.[20] Michod focuses on the evolutionary path that leads from being an autonomous entity to being a part of a larger system. He maintains that the way to model this transition is to focus on the manner in which resources are divided between maintaining oneself and reproducing, the twin goals of viability and fecundity. He argues that due to limited resources these two goals cannot both be maximized. If resources are allocated to viability, then they are subtracted from fecundity and vice-versa. However, if group fitness is understood to be different than the sum (or average) of the fitness values of the individuals, then it can turn out that the group may benefit from division of labor.[21] Some of the cells in a colony can focus on viability and others on fecundity. In an extreme case, when an individual focuses entirely on the viability of the group, its individual's fecundity goes to zero. However, the fitness of the group may still be very high, higher in some cases than the average fitness of the individuals that compose the group. In this situation, when a transition to complete germ-soma specialization occurs, we have the evolutionary emergence of a higher-level entity whose parts have lost their identity as Darwinian individuals, that is, as entities that reproduce. Such Darwinian entities have their own intrinsic ends. The cells of the newly emerged multi-celled organisms have given up such ends to

become parts of something else that does have its own intrinsic ends.[22]  A new Darwinian entity

has emerged at a higher biological level.  In the case studied by Michod and considered by

Okasha, a multi-celled organism has emerged from a group of single-celled organisms.

I propose to use this sort of criterion to understand how individual cognitively agential

entities merge into a genuinely new entity, a WE-ness.[23]  Such an entity is the analogue of the

multi-celled organism.  The analogue of the pursuit of vitality and fecundity is the pursuit of

joint, as opposed to individual, cognitive ends and their proliferation.  The union is accomplished

when cognitive entities give up their own cognitive fecundity to invest in the cognitive vitality of

the emergent multi-cellular cognitive entity.[24]  Their investment in vitality promotes the

fecundity of the emergent cognitive entity, that is, its success in bringing about cognitive

products, a success that can be measured by the reduplication of these products in other cognitive

entities within the same generation and subsequent generations.

We can specify the analogy more precisely using a generalized selection theory account

that involves (1) replicators, (2) units of selection, (3) selected-for capacities and (4) selecting

environments.[25]  For example, suppose the replicators are cognitive traditions.  The units of

selection are the groups that embody these traditions.  Their selected-for capacities are those that

successfully apply these traditions to the problems that the traditions pursue.  And the selecting

environments are the ontological configurations that are the solution sets for the pursued

problems.  Though I have formulated this analysis in terms of purely cognitive ends, I intend its

application more broadly to include ends achieved by broadly cognitive means.  Thus, I have in

mind cultural achievements such as art, law, science, social organization, government etc.[26]

Let us call these group human cognitive achievements cognitive *enhancements*. The WE-ness hypothesis is one possible explanation of this phenomenon. But there are other competing explanations that invoke different sorts of understanding of these group products.

I turn now to findings in developmental psychology indicating that humans' capacity for we-intentionality helps to explain their distinctive cultural and social achievements relative to human primate cousins. I then argue that two related characteristics of we-intentionality, its normative and supra-personal intentionality, provide *prima facie* support for a WE-ness account of cognitive enhancements rather than summative and transcendental social constructionist accounts.[27]

## VI. We-Intentionality and WE-ness

Though some non-human animals are embodied, embedded and extended, making use of tools, engaging in social learning and cultural achievement, researchers agree that they have not developed the kind of cumulative cultural evolution and social institutions typical of the human species.[28] They have not produced science, technology, art, law, governments, nations and so forth. What are the sources for these cognitive enhancements?

Developmental and comparative psychologist Michael Tomasello and his colleagues maintain that a basic source for these achievements is humans' capacity for we-intentionality, an ability to form supra-personal normative intentions.[29] They have found that infants around one year's old have capacities both to help and to share goods and information. The basis of these capacities and what enables humans to develop the cumulative culture and multiple social and cultural products that distinguishes humans from primates is their capacity for "we-

intentionality." This "shared intentionality [is], most basically, the ability to create with others joint intentions and joint commitments in cooperative endeavors."[30]

Further, as infants grow, their caregivers -- through modeling, communication and instruction based on social norms -- shape these innate we-intentionality based capacities.[31] The guiding norms are those of cooperation (including moral norms) and conformity. Moreover, human infants learn in a distinctive fashion because they learn not only by imitation, as do non-human primates, but also by means of caregivers' active teaching. In addition, children not only passively conform to norms (under the influence of adults and their peers) but they also search out what the norms and rules are for a given situation. Indeed they participate in enforcing norms. And, significantly, humans learn from others by a kind of imitation that is motivated merely by the desire to be like them. Conformity learning plays a significant role in the retention of successful joint human practices and such conformity is itself enforced by sanctions and the threat of sanctions.

For instance, three year olds object to a puppet that plays a *one-player* game differently than what the rules allow. They object not merely because the puppet plays the game differently but also because the puppet is not following the rules of the game. When the puppet violates the rules they say things like "One can't do that!"[32] These rules are not merely instrumental or regulative, that is, rules that aid in social interaction, they are constitutive rules -- rules that make the game the sort of game that it is. Tomasello concludes that children view these rules as those of impersonal supra-individual entity.[33]

Tomasello argues that this phenomenon of norm enforcement is based on we-intentionality. The we-intentionality involves an impersonal normativity.[34] We-intentionality reveals itself both in the following of norms of cooperation and in the adherence to norms of

conformity as well as the performance of actions aimed at identifying the actor with a group, initially with the person's significant others, parents, family and schoolmates and then larger cultural groups.[35] We-intentionality then involves and fosters both a kind of group identity and social rationality that operates along with social pressures of various kinds.[36] [37]

Thus, on Tomasello's hypothesis, it is we-intentionality (featuring both normativity and a non-reductive supra-personal intentionality) -- first manifested in the infants innate altruistic capacities for helping, sharing goods and providing information that are themselves selectively shaped by norm-based socialization -- that lies at the basis of human's distinctive social and cultural achievements, among which are the phenomena of cognitive enhancement.


## VIII. Assessing Accounts of Cognitive Enhancement

I assume that competing hypotheses concerning the sources of cognitive enhancements accept that they are distinctively human cognitive achievements. If Tomasello and his colleagues are correct, then these achievements rest on the distinctive human capacity for we-intentionality. We-intentionality displays the features of normative and supra-personal intentionality. I contend that these features of we-intentionality support and are explained better by the WE-ness hypothesis than by either of two competitors, transcendental social constructionism or summativism.[38]

Summative accounts of social phenomena such as Bratman's reduce social phenomena to individual phenomena since they deny the existence of a plural subject, though they allow for joint intentionality.[39] They also do not build the notion of normativity into their accounts of joint intentionality.[40] Thus findings concerning the supra-individuality and normativity of we-intentionality lend support to the WE-ness hypothesis rather than summative accounts.

Antti Saaristo proposes a local sort of socially constituted transcendental constructionism concerning human agency.[41] He argues that social practices are the source of attributions of agency, whether that agency is individual or collective. Examining some findings from social psychology and experimental economics, he suggests that these findings underdetermine competing summative and pluralist accounts of agency and that the reason for this under-determination is not because these disciplines are in the early stages of development. Rather, it is because the attribution of agency, whether individual or collective, is the result of social practices of attributing reasons, an activity distinct from scientific endeavors to find the causes of things.

Saaristo's claims are problematic. First, he examines only a very limited set of scientific findings. Second, his claim that such findings do not support a realistic understanding of social phenomena are based on the bare assertion that the findings concerning either individual or collective agency can be reinterpreted to support the opposite conclusion.[42] Third, his appeal to Hacking's looping effect to explain the construction of individualistic or collective agency is completely compatible with a realist account of group agency, whether, individualistic, summative or pluralist.[43] Fourth, Saaristo's claim that the existence and characteristics of human agency -- whether in its individualistic or collectivist form -- depends constitutively on the interpretive activity imbedded in social practices faces a substantial burden of proof. This is especially so given the social psychological evidence about groups and group identification and the developmental evidence not only about we-intentionality but also about theory of mind and the apparent role that these capacities play in explaining both individual and joint agency.

## VII. Conclusion

I suggest augmenting Andy Clark's 3E-ness thesis that the mind is embodied, embedded and extended, with a WE-ness version of extended cognitive agency, involving groups of human cognitive agents.  I offer a biological model for understanding its ontology.  I then use Michael Tomasello's and colleagues' proposal that humans' capacity for we-intentionality (featuring a normative super-personal intentionality) provides a basis for uniquely human cognitive achievements.  On that basis I suggest that the WE-ness hypothesis is *prima facie* superior to either a summative or transcendental social constructivist account.

---

[1] Notably recently, for instance, Andy Clark, *Supersizing the Mind: Embodiment, Action, and Cognitive Extension* (Oxford: Oxford University Press, 2008), [hereinafter referred to as "Supersizing"] and "Curing Cognitive Hiccups: A Defense of the Extended Mind," *Journal of Philosophy* Volume CIV, Number 4, April, 2007, pp 163-192, (hereinafter referred to as "Curing").  Confer, for instance, Robert A. Wilson and Andy Clark, "How to Situate Cognition: Letting Nature Take its Course" in M. Aydede and P. Robbins (eds.), *The Cambridge Handbook of Situated Cognition* (Cambridge: The Cambridge University Press, 2009), pp. 53-77.

[2] Clark argues for 3Es through an examination and analysis of multiple empirical, hypothetical, imaginative and intuitively possible scenarios.  However, in "Supersizing" and "Curing" he devotes a major portion of his efforts to refuting critics.  He interprets most of his critics to be foes of 3E, especially of embeddedness and extendedness.

[3] I am far from cognizant with the vast literature in this area.  Thus far I have found only one piece in which Clark addresses the issue of the kind of extended mind with which I am concerned (See "Economic Reason: The Interplay of Individual Learning and External Structure" in *The Frontiers of the New Institutional Economics* The Academic Press, 1997).  In other places, I have found that he refers to possible cases of what I would call WE-ness, but he considers such cases to be speculative and does not pursue them.  Among other supporters of 3E-ness, Robert Sutton clearly recognizes the importance of WE-ness and has done work to show how 3E-ness might be applied to extended memory.  See, for instance, his "Exograms and Interdisciplinarity: history, the extended mind, and the civilizing process" in Richard Menary (ed), *The Extended Mind*  (Cambridge, Massachusetts: MIT, 2010).

[4] This is not to mention historical philosophical precedents in Idealism and Personalism, as well as in Indian religious traditions, pantheism and panentheism.

[5] For a very helpful survey see Deborah Tollefson, "Collective Intentionality," *The Internet Encyclopedia of Philosophy*: http://www.iep.utmedu/coll-int/.

[6] More precisely, the notebook embeds dispositional beliefs that fulfill such criteria as reliability, portability and accessibility, allowing it to serve to some extent the role that Otto's failing memory once served.

[7] Far from being a paradigm example of a super-sized mind, the original thought experiment points to a relatively minimal sort of embodied, embedded and extended cognition. Moreover, amazingly, Otto is presented in Clark and Chalmers' account merely as an example of a reasonably postulated version of the extended mind hypothesis. Though I have not yet been able to examine Clark's extensive work on embodied, embedded and extended cognition, I have found that he has a more expansive and less tentative views in earlier writings. See, for instance, *Being There*: Putting Brain, Body, and World Together Again (Cambridge Massachusetts: The MIT Press, 1997) and an essay on economics and economic systems "Economic Reason: The Interplay of Individual Learning and External Structure" in *The Frontiers of the New Institutional Economics* The Academic Press, 1997. However, in a recent work, he and Rob Wilson continue the modest approach of the original thought experiment. (Robert A. Wilson and Andy Clark, "How to Situate Cognition: Letting Nature Take its Course" in M. Aydede and P. Robbins (eds.), *The Cambridge Handbook of Situated Cognition*.) As far as I have been able to discern in his latest volume ("Supersizing"), Clark never alludes to such unities as examples of 3E-ness, except in the Appendix to this volume, containing an earlier essay co-authored with David Chalmers. In the Appendix to the volume under discussion, in a paper co-authored by Clark and Chalmers, there is an allusion to social unities (*Supersizing*, Pages 231-32). There they *speculate* about the applicability of 3E-ness to "socially extended cognition". Thus, they comment: "What about socially extended cognition? Could my mental states be partly constituted by states of other thinkers? We see no reason why not, in principle. In an unusually interdependent couple, it is entirely possible that one partner's beliefs will play the same sort of role for the other as the notebook plays for Otto. What is central is a high degree of trust, reliance, and accessibility. In other social relationships these criteria may not be so clearly fulfilled, but they might nevertheless be fulfilled in specific domains. For example, the waiter at my favorite restaurant might act as a repository for my beliefs about my favorite meals (this might even be construed as a case of extended desire). In other cases, one's beliefs might be embodied in one's secretary, one' accountant, or one's collaborator." (*Supersizing*, Pages, 231-32) I note that these examples of social extension are limited to the perspective of one cognitive agent using another cognitive agent as an instrument. Moreover, the authors treat them as speculative extensions. My aim is to show that an extended form of 3E-ness, WE-ness is scientifically theoretically plausible and empirically supported.

[8] Of course, Clark and Chalmers's story is entirely compatible with my back-story. The point of the back-story is to illustrate a richer notion of embodied, embedded and extended cognitive agency, in particular WE-ness based cognitive agency. I also have an implicit methodological point. Thought experiments take us only so far. In a sense they come a dime a dozen and even more cheaply for those with creative imaginations. Although I shall not argue for it here, my position is that metaphysical claims, at least those about this world, are best supported by our best current relevant scientific theories and results. Thus I shall attempt to support my claims about We-ness with well-supported empirical findings and scientific theories.

[9] I do not in this paper pretend to engage the highly developed analytic literature in social metaphysics. For a very helpful recent collection of essays, see Frederick F. Schmitt (ed.) *Socializing Metaphysics: The Nature of Social Reality* (Lanham, Maryland: Rowman and Littlefield) 2003. Schmitt's introductory chapter is very helpful. My approach is to sketch a metaphysics for WE-ness based in currently accepted theories and findings of the relevant sciences. This metaphysics is, therefore, a restricted one that concerns the social reality of this world, not all possible worlds. Thus it is neither an exercise in the analysis of "our concept" of the social nor an attempt through the method of cases to determine necessary and sufficient conditions for what holds about social concepts in all possible worlds or what must be the case about the phenomenon in all possible worlds. Moreover, as with scientific theories generally, the findings concerning the nature of social phenomena will at least to some extent under-determine the competing theories. Thus, the brief conclusions that I draw at the end of this paper about the superiority of the WE-ness account of social phenomena relative to its competitors are highly tentative.

[10] Although I cannot develop it here, I believe that an organizational account of biological functions, rather than dispositional or etiological accounts, offers the best way to understand in a scientific fashion, the notion of function. See Matteo Mossio, Critian Saborido and Alvaro Moreno, "An Organizational Account of Biological Function," *British Journal of the Philosophy of Science* (60) 2009: 813-841.

[11] Godfrey-Smith, Peter, *Complexity and the Function of Mind in Nature*. Cambridge: Cambridge University Press, 1996.

[12] Although I cannot argue for it here, I contend that the ontological kinds that I propose are scientifically plausible. In particular, it should be noted that I assume a working notion of intrinsic ends and cognitive kinds. Much of the debate between proponents and opponents of the extended mind thesis centers on questions about nature of cognition. In his recent paper "From Cognition's Location to the Epistemology of it Nature," Matthew Barker argues that the opposing theories of extended versus merely embedded mind are empirically undecidable. He maintains that this fact should lead us to abandon trying to understand *where* cognition is located and find out first *what* cognition is. This is an eminently reasonable suggestion. However, it turns out that in Barker's presentation of the problem, the key issue concerns the differences between the constitution relation and the causal relation, not as far as I can see, fundamental differences between the parties concerning the nature of cognition. In this paper, I assume an understanding of the nature of cognition that is, I think, acceptable to all parties. I believe that my attempt to distinguish criteria for entity-hood will be useful in coming to the genuine issue, identified by Barker, concerning the differences between constitution and causation, though I cannot argue for that claim here. In her recent "Neuroscience and the Multiple Realization of Cognitive Functions," (Philosophy of Science (77) 2010, 419-456), Carrie Figdor suggests a biologically-based criterion (degeneracy) for distinguishing cognitive systems that I believe is extremely relevant for ascertaining how neuroscientists actually solve issues concerning constitution and causation. Barker, I believe, addresses the differentiation between constitution and causality at a level of abstraction that makes their empirical testability problematic.

[13] Though not making the case for it here, I contend that such cognitive and agential unities are plentiful and growing. Clark starts at the lowest level of cognitive engagement and argues against opponents of even minimal supersizing. He focuses on embodied, embedded and extended cognition where the embedding environment is non-living artifacts or natural things. As a consequence the extended cognition that he reflects on concerns cognitive unities of humans and these sorts of non-living artifacts or natural things. I start at the other end of the continuum with a super-sized mind of the We-ness sort, that is, where the unities concern groups of human cognitive agents. My concern is to analyze, understand and explain these multiple cases this sort of 3E-ness with which we seem to be confronted. Using a more familiar sort of distinction, I shall maintain that the type of super-sizing exhibited in the WE-ness form of 3E-ness occurs when the agents are intentional agents and the roles they play are such that none is used as a mere instrument.

To see the contrast, consider the limits of the Otto case. Clark and Chalmers urge us to notice that the epistemic brain is embodied by means of sense organs and embedded by epistemic tools -- things that have no intrinsic ends, but can by artifice be fashioned into epistemic tools, that are at hand in the environment. They then point to these embodied and embedded brains and note how they form dynamic epistemic unities. Clark and Chalmers' 3Es start off as human individuals and remain human individuals even as they form epistemic alliances with bodily and non-bodily, non-living parts. I start with recognized cognitive agents and explore how they might form a supra-individualistic cognitive agent.

[14] Richard Dawkins, *The Selfish Gene* (New York: Oxford University Press, 1976.)

[15] Clark explicitly compares his extended mind thesis to Dawkins' gene centered view in "Curing Cognitive Hiccups: A Defense of the Extended Mind," *Journal of Philosophy* Volume CIV, Number 4, April, 2007, pp 163-192.

[16] I shall not here go into the detailed discussions of the relative merits of gene, organism and group-centered approaches to questions about the units of selection. See, for instance, Sterelny, Kim and Paul E.

Griffiths, *Sex and Death: An Introduction to Philosophy of Biology* (Chicago: The University of Chicago Press, 1999).

[17] Humans come with epistemic embedding capacities and learn to develop even more of them. Confer Kim Sterelny, *Thought in a Hostile World: The Evolution of Human Cognition* (Malden, MA: Blackwell Publishers, 2003).

[18] J. Maynard Smith and E. Szathmary, *The Major Transitions in Evolution* (Oxford: Freeman, 1995). Confer also Elliott Sober and David Sloan Wilson, *Unto Others: the Evolution and Psychology of Unselfish Behavior* (Cambridge, MA: Harvard University Press, 1998).

[19] Samir Okasha, "Individuals, groups, fitness and utility: multi-level selection meets social choice theory." *Biology and Philosophy* (2009) 24, 561-564. He makes use of social choice theory as a heuristic for deriving this criterion. I will not discuss how he does that but rather turn to his account of the criterion for the emergence of levels that are then subject to multi-level selection.

[20] Here he follows the work of Michod on volvocine green algae. Confer Rick Michod, "The Group Covariance Effect and Fitness Trade-offs during Evolutionary Transitions in Individuality." *Proceedings of the National Academy of Science* 103 24 (2005), 9113-9117: R. Michod and A Nedelcu "On the Reorganization of Fitness during Evolutionary Transitions in Individuality." *Integrative Comparative Biology* 43 (2003), 64-73 and R. Michod, Y. Viossat, C. Solari, M. Hurand and A. Nedelcu "Life History Evolutionary and the Origin of Multicellularity," *Journal of Theoretical Biology* 239 (2006), 257-72.

[21] In social choice theory, the analogue of group fitness, social utility is often considered to be the sum of or the average of individual utilities.

[22] The cells of various types still have functions and in this sense have intrinsic ends. But these ends are ordered in normal, well-functioning organisms to the higher-level intrinsic ends of the organism. They also reproduce themselves through cell division, but their self-maintenance and reproduction is also ordered to that of the self-maintenance and reproduction of the organism of which they are living parts.

[23] In speaking of cognitive ends and cognitive agents, I do not have in mind merely cognitive activities such as that engaged in by a scientific research team. I intend all those activities, whether their goals be cognitive products or not, that are performed by agents using their representational capacities.

[24] Normative scientific practices are indicative of such entities. See, for instance, Melinda B. Fagan, "Social Constructionism Revisited: Epistemology and Scientific Practice," *Philosophy of Science* 77, #1, 2010, pp. 92-116. The structure of successful cognitive groups is likely to be quite complex. See, for instance, Michael Weisberg and Ryan Muldoon "Epistemic Landscapes and the Division of Cognitive Labor," *Philosophy of Science* (76) 2009, pp. 225-252.

[25] Lindley Darden, L. and Joseph Cain, "Selection type theories," *Philosophy of Science* 56 (1989). 106–29.

[26] I believe the analysis will also fit smaller scale informal groupings, though I shall not attempt to show that here.

[27] There are echoes of the theory of extended minds and WE-ness in analytic social ontology. The analytic tradition provides a rich and varied set of analyses of plural subjects (For very helpful surveys see Frederick F. Schmitt (Ed.) Socializing Metaphysics: The Nature of Social Reality (Lanham, Maryland: Rowman and Littlefield, 2003) and Deborah Tollefson, "Collective Intentionality," *The Internet Encyclopedia of Philosophy*: http://www.iep.utmedu/coll-int/). Many philosophers treat these analyses as analyses of our *conception* of plural subjects or of collective intentionality or belief; but they can also be

considered as attempts to capture the phenomena in question by means of a theoretical hypothesis that requires empirical support. I shall treat them in that fashion.

[28] Though, as far as I can ascertain Clark does not do so, it seems perfectly consistent with his ideas of embodiment, embeddedness and extendedness, to extend these concepts to the non-human animal activities. The case for extended non-human action appears plausible. So too does an analogous case for non-human WE-ness. In these cases non-human individuals would form a group that itself constitutes a supra-individualistic entity that produces enhancements beyond those that an individual can accomplish. Confer, for instance, Brian Skyrms' comments on Tomasello's views in the Forum Section, of Michael Tomasello, *Why We Cooperate* (Cambridge: The MIT Press, 2009), pp. 137-148.

[29] Michael Tomasello, Malinda Carpenter, Josep Call, Tanya Behne, and Henrike Moll, "Understanding and sharing intentions: the origins of cultural cognition, *Behavior and Brain Sciences* (2005) 28, 675-735 and Michael Tomasello, *op. cit*. Of course, Tomasello and colleagues hypotheses do not go uncontested. Both the above-mentioned article and Tomasello's book are accompanied by detailed critical comment. I do have the space here to examine and assess the merits of these criticisms.

I note also that work by social psychologists, anthropologists and others on both the role of groups and normativity in human thought and activity is compatible with Tomasello and colleagues findings about we-intentionality. For instance, social psychologists consider psychological features that suggest non-reductive intentionality and that seem to play a role in the formation of a group. (Confer, for instance, David G. Meyers, *Social Psychology*. Boston: McGraw Hill, 2008.). They find that (1) we categorize in terms of human groups, placing others and ourselves into groups, (2) we identify ourselves with certain groups (in-groups) and others with different groups (out-groups), and (3) that we contrast our group with out-groups, giving a favorable bias toward our own group. In addition, (4) such group bias leads to preferential liking of our own and dislike for out-groups. Further, (5) group membership becomes a part of self-identification that enables and leads to the following of group norms and (6) group success leads to feelings of individual success. Moreover, (7) group membership also involves relative status within a group and to attitudes of superiority and inferiority. And (8) in-group bias and relative status lead to and supports prejudicial attitudes and (9) the latter makes for self-fulfilling prophecies as well as self-perpetuating stereotypes. Thus, we find what Hacking (The Social Construction of What? Cambridge, Mass.: Harvard University Press, 1999) and others have described as the social construction of social reality.

What about normativity? I have suggested that a working criterion of We-ness is a feature of the end directedness of its parts. The parts take up as their goal, the goal of the whole. In the case of cognitive agents that goal is cognitively shared and acted upon. The workings of the parts are dictated by the goals of the whole. They become norm directed in the sense that the goals and sub-goals dictate means that can be understood as the norms under which the group operates.

One of the striking features of human life is its extensively normative character. Sripada and Stich [Chandra Sekhar Sripada and Stephen Stich "A Framework for the Psychology of Norms," in Peter Carruthers, Stephen Laurence and Stephen Stich (Eds.) *The Innate Mind*, Volume 2: Culture and Cognition (Oxford: Oxford University Press, 2006), pp. 280-301.] take a norm to be a rule or principle that specifies actions that are required, permissible or forbidden independently of any legal or social institution." (Ibid. p. 281) As such they provide intrinsic motivation for action. They offer this notion of norms not as a piece of conceptual analysis of our ordinary conception of a norm or as a technical analytic definition. Rather they consider it to be a theoretical hypothesis about a natural kind that plays a significant role in the social sciences. They maintain that this hypothesis is based on well-supported empirical generalizations about both social level and individual level facts found in the social sciences.

On the social level norms are found to be (1) cultural universals and (2) very ancient. For these reasons, (3) they are likely to be implemented by innate psychological mechanisms. In addition, (4) they play a major role in the lives of peoples in all cultures. But (5) the contents of the norms of different groups vary widely and (6) these content variations are much greater between groups than within groups where the norms tend to be homogenous. Nevertheless, (7) these norms have a degree of cross cultural similarity in so far as they concern certain broad areas of action including, for instance, prohibitions on

killing, physical assault and incest. But, within these broad similarities, there are wide variations. Moreover, (8) the generalizations about what is prohibited, required or allowed have exceptions.

On the individual level, it has been found that (1) norms have a reliable pattern of ontogenesis, (2) individuals acquire some norms fairly early in life, (3) as well as some capacity to reason about norms. Researchers have also found that (4) members of cultures acquire their respective cultural norm variation by the age of nine and retain it thereafter into adulthood. In addition, (5) norms are powerfully motivating, to such an extent that despite the existence of other motivations that sometimes overwhelm them, norms are intrinsically motivating. This is displayed in the sometimes-unselfish character of the actions that they inspire and the fact that they are internalized thus enabling the kind of reliable compliance to their injunctions from early life that members of a culture generally display. Sripada and Stich present significant empirical support for the phenomenon of intrinsic motivation from both social psychological research and experimental economics. These include the well-known research of Bateson on empathy-based helping and in experimental economics on third-party punishment for norm violations.

Finally as an aid in the explanation of these social and individual phenomena about norms, Sripada and Stich offer a general sketch of the structure (architecture) of the psychological mechanisms involved in norm governed activities and practices. They quite reasonably suggest that such a structure will contain acquisition and execution mechanisms. These on first pass would involve on the acquisition side a capacity to identity norm implicating behavior on the basis of proximal environmental clues as well as a capacity to infer the content of these norms. On the execution side are norm databases and rule-related reasoning capacities as well as motivations to comply with rules and to punish non-compliance. They go on the fill out the execution architecture in terms of emotions and reasoning. Sripada and Stich recognize that what they are offering is only a sketch that needs to be filled in with empirical findings and well-supported theoretical accounts of acquisition and execution mechanisms. Such filling in is, of course, a matter of continuing endeavor. The empirical findings and a theoretical proposal based on these findings being developed by developmental psychologist Michael Tomasello and his colleagues concerning we-intentionality are one important way of understanding these acquisition and execution devices.

[30] Michael Tomasello, *op. cit.*, p. xiii. These abilities are supported by the capacities for joint attention and mutual knowledge and the motivation to help and share with others. Tomasello and colleagues provide significant empirical support for the claim that beginning at about one year of age infants spontaneously display various sorts of helping behaviors. (Of course, they also display self-centered behaviors.) They argue that this capacity for helping is in-built and shaped by caregivers as the infant grows to fit the social norms of the group. The helping behavior is displayed in sharing of goods and information and in the providing of services.

[31] Tomasello argues that these capacities are innate. Evidence for their innateness is based on these features: their early emergence, immunity from encouragement and undermining by rewards, deep evolutionary roots in great apes, cross cultural robustness, and their foundation in natural sympathetic emotions. Tomasselo (op. cit., p. 13) gives these reasons with respect to the capacity to help. But later he extends the innateness claim to sharing goods and information (*Op. cit.*, p. 28). See also, Tomasello et al, *op. cit.*, pp. 675-735.

[32] Tomasella, *op. cit.*,, p. 37 and note 36.

[33] "This shows that children view even simple conventional norms of how a game is played not just as instrumental guides to their own effective action – actions likely to please powerful adults or garner some other reward – but as supra-individual entities that carry social force independent of such instrumental consideration" (Tomasella, *op. cit.*, pp. 37-38). Moreover, in learning the game and its rules the child needed only the demonstration of how to play the game. She did not also need to see the adult make a mistake and correct herself. Nor did the adult need to use any normative judgments or language before the children "jumped to normative conclusions about how the game should be played" (Tomasella, *op. cit.*, p. 38).

[34] "My proposal, therefore, is that children's respect for social norms is not due solely to their sensitivity to authority and reciprocity. From a young age, children also possess a kind of social rationality along the lines of what the philosopher Thomas Nagel proposes in *The Possibility of Altruism*, what we might call a 'he is me' attitude of identification with others and a conception of self as one among many, leading to the impersonal ' view from nowhere'." (Tomasella, *op. cit.*, p. 40)

[35] Tomasello maintains that these results are in accord with G. H. Mead's theories concerning identification with significant others and with Mead's generalized other.

[36] Internalization of these norms, as well as the emotions of guilt and shame, enhances the following of both norms of cooperation and norms of conformity. In addition, we-intentionality leads infants to enforce norms not only on themselves through guilt and shame but also on others. It seems then that the phenomenon of we-intentionality is a good candidate for the psychological acquisition and execution structure for norm possession postulated by Sripada and Stich.

[37] Tomasello and his colleagues propose a hypothesis about the evolutionary origins of human altruistic capacities and cooperativeness (in the larger sense of humans' tendency and ability to live and operate together in institution-based collective groups) by comparing humans with their nearest primate relatives.

[38] Though I cannot here go into the details of the analyses deriving from analytic social ontology, outlining the different sorts of accounts will help us discern the import of the findings of concerning we-intentionality that I want to discuss.

Social constructivist maintain that the attribution of either collective or individual intentionality, either we or I attributions, or for that matter the attribution of any sort of agency, is itself a matter of social practices that provide *reasons* for behavior as opposed to assigning *causes* of behavior. On this view reasons are distinct from causes and thus the giving of reasons is a social practice quite distinct from that of determining the causes of a behavior. The former is a practice that constructs agency and its individual and collective forms.

On the other hand, theorists that maintain that some reasons are causes understand that psychological states that are intentional play a role in the bringing about of actions. Among these theorists we find instrumentalist and realist accounts of collective intentionality. The former views our ordinary talk about what *we* intend and believe, for instance, as well social scientific theories concerning collective agents and their intentional states as helpful fictions. The latter understands such theories and ordinary usage realistically. We can divide the realists accounts into summative and non-summative ones. John Searle and Michael Bratman offer summative accounts while Margaret Gilbert and RaimoTuomela present non-summative views. The former, as the name implies, reduces collective agency and intentionality to the summation of the agency and intentionality of individual agents, while the latter argues that the phenomena are emergent, not reducible to the sum of its parts. Besides reductivism and emergentism, there is eliminativism advocated by, for instance Frederick Schmitt. And to round things out some philosophers, for instance, Peter Winch, argue that the concept of the social is basic and that an understanding of the individual is dependent on an understanding of the social.

Among the emergentists, Margaret Gilbert focuses on small informal groups and Raimo Toumela on both these and large formal groups such as corporations. The non-summative accounts are more ontologically expansive, but I believe that the empirical evidence provides support for it

Consider Gilbert's view of what counts as a plural subject of some action or psychological state. Gilbert takes her account to provide the necessary and sufficient conditions for the concept of plural subject. As I mentioned, I will view such accounts as theoretical hypotheses that need empirical support. "Individuals A1…AN form a plural subject for Xing (for some action X or psychological attribute X) if and only if A1…AN form a joint commitment to Xing as a body. A joint commitment to act as a body is a commitment made by a collection of individuals to perform some present or future action as would a single individual. Joint commitments are formed when each of a number of people expresses his or her willingness to participate in the relevant joint commitment with the others" (Tollefson, Ibid.). That obligation is special because (1) it allows for coercion, (2) cannot be rescinded by an individual member,

(3) provides a reason to act and (4) is something of which the participants are aware.  Each individual understands that the joint commitment forms only when all the relevant people have agreed to participate.

Tuomela analyses groups and intentionality where such joint agreement may not be present in all the members of the group, making a distinction between operative and non-operative members, for instance, in large corporations like British Petroleum.  On Gilbert's view, these agreements establish an obligation of the members to each other.  Though Gilbert's and Tuomela's views are significantly different, they share some features that distinguish them from the summative views and upon which I want to focus as I examine the developmental evidence.  These are (1) non-reductive or supra-personal intentionality and (2) normativity.

[39] Confer, Tollefson (*op. cit.*,) and Michael Bratman, *Intentions, Plans, and Practical Reason.*  (Stanford, California: CSLI Publications, 1999).

[40] Bratman recognizes the normativity sometimes attaches to we-intentionality, but does not make it an essential feature of we-intentionality.  (As discussed in Tollefson, *op. cit.)*

[41] Antti Saaristo "There Is No Escape from Philosophy: Collective Intentionality and Empirical Social Science," *Philosophy of the Social Sciences*, Volume 36, Number 1, March 2006, pp. 40-66.

[42] Such reinterpretations are clearly always logically possible.  The question is whether, given both common sense and scientific findings, such reinterpretations are plausible, in particular, as plausible, as summative or pluralistic interpretations.  I don't find them to be so.

[43] Realists can grant that some overt social realities do require this sort of constitutive dependence, for instance, Searle's cocktail party case.  Ron Mallon (Naturalistic Approaches to Social Construction, *Stanford Encyclopedia of Philosophy*, 4/23/10; "A Field Guide to Social Construction," *Philosophy Compass*, Volume 2, Issue 1, pp. 93-108 and "Social Construction, Social Roles, and Stability," in Frederick F. Schmitt (Ed.) Socializing Metaphysics: The Nature of Social Reality (Lanham, Maryland: Rowman and Littlefield, 2003), pp. 327-354) makes a similar sort of point in discussing constitutive constructionist claims.