TESIS DOCTORAL

Culture as a collective mind.

The emergent properties of a community

DOCTORANDO

Francesco Consiglio

DIRECTOR

Dr. Fernando Martínez Manrique

Dpto. de Filosofía I

Programa de Doctorado en Filosofía

UNIVERSIDAD DE GRANADA

GRANADA
2020

Editor: Universidad de Granada. Tesis Doctorales
Autor: Francesco Consiglio
ISBN: 978-84-1306-538-0
URI: http://hdl.handle.net/10481/62929

Publications

Consiglio, Francesco (forthcoming) «An embodied mind. Merleau-Ponty's influence on contemporary neuroscience», in *Ontology an ICT in Education*. Conference proceedings.

Consiglio, Francesco (2019). «Collective intelligence and social ontology. Bridging the divide between human and animal collective cognition through stigmergy and Peircean semiotics». *Disputatio. Philosophical Research Bulletin* 8, no. 11: pp. 531–547.

Consiglio, Francesco (2018). «Collective intelligence: An emergent semiotic system». *Disputatio. Philosophical Research Bulletin* 7, no. 8: a009.

Consiglio, Francesco (2017) «Intelligenza di gruppo e sistemi multi-agente». *Quaderni della Ginestra* 22:3, pp. 72–85.

Consiglio, Francesco (2017). «Information networks and systemic properties. An epistemological perspective». *Disputatio. Philosophical Research Bulletin* 6:7, pp. 309–321.

Ai miei genitori, a Clara

AGRADECIMIENTOS / ACKNOWLEDGEMENTS / RINGRAZIAMENTI

Esta tesis trata de la cultura como mente colectiva y colectiva es, en un profundo sentido, toda empresa de investigación. Quiero por lo tanto agradecer a todas las personas que, cada una en su medida, han influido en este trabajo.

Quiero en primer lugar agradecer a mi director, Fernando Martínez, por haber creído en mi tesis cuando aún era un pequeño borrador, con ideas tal vez sugerentes pero aún verdes. Quiero agradecerle el gran trabajo que ha conllevado dirigirme durante estos años en Granada, todas sus sugerencias, que me han orientado de manera determinante en el desarrollo de mis ideas. Quiero agradecerle las muchas discusiones y la atenta lectura de todos mis borradores. Quiero agradecerle, por supuesto, su enorme paciencia y, en general, el gran ejemplo que me ha estado dando como investigador y como persona.

I would also like to thank Tillmann Vierkant who was my supervisor during my research stay at the School of Philosophy, Psychology and Language Sciences of the University of Edinburgh. He red my drafts and kindly gave me very insightful suggestions.

Quiero darle las gracias a todos los miembros del Departamento de Filosofía I de la Universidad de Granada, que me han acogido en un ambiente filosóficamente muy activo. Particularmente, quiero agradecer a Manuel de Pinedo y Neftalí Villanueva los muchos comentarios costructivos que me han hecho en las ocasiones en las que he podido presentar mis investigaciones en el departamento. A Manuel de Pinedo le debo la introducción a los estudios sobre Teoría de Nicho, un tema que tiene mucha relevancia en esta tesis.

Quiero agradecer especialmente a Víctor Fernández Castro por sus comentarios acerca de la cognición de grupo y a Manuel Heras Escribano por los suyos acerca de la cognición ecológica.

A María José García Encinas y a Javier Cumpa quiero darles las gracias por sus comentarios sobre los últimos desarrollos de esta investigación, especialmente por lo que atañe el concepto de categoría.

Quiero agradecer a mis compañeros de doctorado, David Bordonaba, José Ramón Torices, Eduardo Pérez, Andrés Soria, Llanos Navarro, Alba Moreno, Manuel Almagro, Javier Osorio, Amalia Haro y Laura Marcon, por la amistad y los comentarios durante los seminarios en estos años. Quiero especialmente agradecer a Nemesio García-Carril Puy, a Pedro Jorge García, a Mirco Sambrotta y a José Ferrer de Luna por todos los grandes momentos, filosóficos y no filosóficos, que me han dado en estos años en Granada. Las palabras nunca pagan la amistad.

Chiaramente, devo tantissimo alla mia famiglia, della quale non mi è mai mancato il sostegno. Grazie ai miei genitori, Eugenio e Rosa, che mi hanno sempre appoggiato in tutte le mie scelte, dovunque mi portassero, aiutandomi a realizzare i miei progetti; grazie alle mie sorelle, Stefania e Laura, che mi hanno sempre saputo ascoltare e aiutare, che non mi hanno mai lasciato solo con le mie difficoltà. Una parte importante di questa tesi è merito vostro.

A Clara, per questi anni insieme, per la pazienza infinita con questo italiano ostinato, per tutte le lunghe e appassionate discussioni sulle idee centrali di questa tesi, per tutti i viaggi e le esperienze, per tutto ciò che mi hai insegnato finora e mi insegnerai in futuro, per tutto il tuo amore e il tuo appoggio, grazie. Senza di te non avrei potuto pensare molte delle cose che penso. *Qué alegría, vivir / sintiéndose vivido*.

CONTENTS

Resumen

Este trabajo surge de una pregunta fundamental: ¿pueden los grupos de personas formar algún tipo de mente colectiva? Mi respuesta es que sí. El principal objetivo de esta tesis, por ende, consiste en defender esta posición, es decir, que las entidades colectivas como los grupos humanos, como las comunidades, tienen sus propias mentes.

Desde luego, hablar de 'mente' en términos generales significa adentrarse en un terreno resbaladizo, ya que 'mente' se dice de muchas maneras. Por lo tanto, he optado por una de ellas: he decidido compromenterme con un concepto de mente considerada como la estructura de información emergente que guía la conducta de un sujeto, organizando sus actividades cognitivas. Teniendo en cuenta que el sujeto en el que me enfoco en esta investigación es una comunidad humana, me ha parecido sensato concluir que aquella estructura de información que emerge dentre de un grupo humano y guía su conducta es ciertamente su propia *cultura*.

Contrariamente a la noción de mente como característica interna del cerebro que ha caracterizado la perspectiva cognitivista contemporánea (Minsky 1988), en esta tesis me comprometo con una perspectiva externista. El marco externista que desarrollo es aquel de un *externismo activo*: es decir, el entorno tiene un papel activo en nuestras actividades cognitivas.

El núcleo de la tesis es el siguiente: me compromento con una noción de cognición entendida como *producción de significado* (Gallagher 2013), que es el resultado de una actividad necesariamente colectiva; defiendo que los significados disponibles para cierta comunidad son manipulados por sus miembros a través de las *representaciones* en las que cristalizan; argumento que tales representaciones, disponibles en un entorno compartido, implican el desarrollo de prácticas normativas que regulan sus usos públicos (Menary 2013); argumento, por lo tanto, que la manipulación pública de representaciones acaba en la emergencia de *cánones representacionales*; defiendo que tales representaciones canónicas emergentes han de considerarse como *creencias colectivas* y han de atribuirse, por ende, a la comunidad entendida como sujeto colectivo; estos cánones, finalmente, retroalimentan a los miembros de la comunidad, moldeando sus mentes, sus categorías y

los propios significados que estos agentes usan en sus actos de referencia. Por esta razón acabo concluyendo que la propia *semántica* es una propiedad emergente de los colectivos. Por lo tanto, en los *nichos culturales* (Laland and O'Brien 2011), o sea, los entornos culturalmente moldeados habitados por comunidades humanas, se dan las condiciones para la emergencia de una *mente colectiva*.

En detalle, partiendo desde un análisis de la Tesis de la Mente Extendida (Clark and Chalmers 1998), presento críticamente el estado de la cuestión acerca del externismo activo (capítulo I) : el argumento clásico de Clark y Chalmers y las críticas de sus detractores se enfocan en la metafísica de la mente; es decir, a favor o en contra de un marco funcionalista, donde se defiende o se ataca el llamado *principio de paridad* entre dinámicas cognitivas internas y dinámicas cognitivas implementadas externamente por medio de artefactos. El objetivo original de Clark y Chalmers es tratar de demostrar que los artefactos externos que suportan nuestros procesos cognitivos, son *constituyentes* de los mismos: por ende, la mente se extiende en tales artefactos cognitivos. Explico que una segunda ola de teóricos de la mente extendida apuesta, al revés, para el llamado principio de complementariedad (es decir, los recursos externos no constituyen sino complementan nuestros procesos cognitivos). Finalmente, tras un análisis detallado de las posiciones típicas de la segunda ola, me enfoco en la llamada tercera ola de la mente extendida. Ésta está caracterizada por una lectura de la mente en cuanto fenómeno social y pone el acento sobre la producción de significado como rasgo característico de la cognición.

En el segundo capítulo delimito la noción de mente colectiva, marcando las diferencias con nociones cercanas, como la de inteligencia colectiva y auto-organización. Clarifíco que hay dos perspectivas principales sobre el concepto de mente que desembocan en dos visiones distintas de la mente colectiva: la perspectiva a la que llamo 'Cartesiana' toma como término de referencia las características de la psicología individual para tratar de detectar características equivalentes a nivel colectivo; al contrario, la perspectiva informacional considera que el rasgo característico de la mente consiste en la manipulación de información, por ende, los autores que apoyan esta perspectiva remarcan la relevancia de las representaciones públicas como constituyentes

de las mentes colectivas. Yo me compromento con la perspectiva informacional y avanzo una hipótesis para la mente colectiva, explicando que la estructura cultural de una comunidad constituye su mente colectiva. Defiendo que las comunidades producen memorias colectivas y creencias colectivas por medio de la manipulación de representaciones públicas en un entorno ecológico.

Seguidamente (capítulo III), explico qué es un entorno ecológico introduciendo la Teoría de la Construcción del Nicho. Se trata de un paradigma que pone en el centro la retroalimentación recíproca que se establece entre el entorno (el medio físico) y los agentes que lo habitan. Destaco, que el entorno tiene un papel crucial en las dinámicas de los grupos humanos que allí viven porque en él estos agentes construyen su nicho cultural. Concluyo argumentando que el nicho cultural de un grupo humano constituye su andamiaje cognitivo.

El cuarto capítulo extiende el análisis del nicho cultural concebido como andamiaje cognitivo. Aquí me enfoco en el uso de artefactos cognitivos dentro del nicho y en la construcción de una memoria colectiva toda alrededor de la comunidad: registros de memoria externos (exogramas) están estructurados en el andamiaje cognitivo de la comunidad por medio de narrativas públicas.

Luego (capítulo V) analizo los distintos tipos de información disponibles en el entorno ecológico humano: remarco la diferencia entre la información ecológica que aprovechamos en forma de *affordances* y, por otro lado, la información representacional que caracteriza el nicho cultural y experimentamos en forma de constricciones perceptivas y convenciones. Aquí detallo el mecanismo de la estigmergia, entendido como la dinámica fundamental de la retroacción dentro del nicho que causa la emergencia de todas las estucturas del nicho mismo. Concluyo el capítulo enfocándome en la emergencia de cánones representacionales dentro del nicho cultural y argumentando que dichos cánones constituyen creencias colectivas, cuyo sujeto es la comunidad misma que las produce. Defiendo que dichos cánones desempeñan un papel cognitivo dentro de la mente colectiva de la comunidad, mientras que, con respeto a los agentes individuales que forman parte de la misma comunidad, tales cánones

desempeñan un papel meta-cognitivo: es decir, regulan sus rutinas cognitivas por medio de un acción de *mindshaping* (Mameli 2001; McGeer 2007; Zawidzki 2008, 2013).

Finalmente (capítulo VI), argumento que la emergencia de cánones representacionales dentro de una comunidad produce lo que llamo 'categorial mindshaping'. Es decir, argumento que los cánones representacionales ejercen una presión normativa sobre la comunidad de usuarios y moldean sus categorías y sus referentes. Concluyo remarcando que la semántica resulta ser una propiedad emergente de los colectivos.

INTRODUCTION

Culture as a collective mind

This work originates from a fundamental question: do groups of people form something like a collective mind? My answer is, yes. The main aim of this thesis, then, will consist in arguing for this point, namely, that collective entities like human groups, like communities, have minds of their own.

Certainly, speaking of 'mind' in general terms is like walking on a slippery surface, because 'mind' is said in many ways. Therefore, I opted for one of them: I decided to commit to a concept of mind regarded as an emergent structure of information which guides the behaviour of a subject, organizing her cognitive activities. Considering that the subject I focus on in this work is a human community, it made sense to me concluding that the structure of information which emerges within a human group and guides its behaviour is indeed its very *culture*. Then, I started to prepare the ground for my argument to defend that the culture of a human community ultimately constitutes its collective mind.

It is always difficult to think of ordinary things in new ways. This is also the case of the mind. We generally use the term 'mind' when we refer to some particular cognitive properties that we humans feature, and we generally think of cognition as something proper of individual subjects: cognition is something that brains do, so there is a brain for each mind. In effect, Marvin Minsky used to say that the mind is what the brain does (Minsky 1988). It is for this reason that the mind has generally been conceived, in the contemporary cognitivist account, as a brain-bounded feature of some organisms. The perspective offered in this work is quite different indeed.

In 1998 two philosophers, Andy Clark and David Chalmers, came along with a disruptive question: «Where does the mind stop and the rest of the world begin?» (Clark and Chalmers 1998, p. 7). Considering that the environment plays a main active role in driving our cognitive processes, they suggested that the mind itself *extends* into the

environment, beyond the narrow limits of skull and skin. This was the first spark which set down the debate about the so called Extended Mind Thesis.

Clark and Chalmers's argument was based on a functionalist approach to this issue: appealing at the parity of functions carried out by both some internal cognitive processes and some external devices, they argued that cognition extends into such devices and so does the mind itself. For this reason, the bounds of the mind turned out to be quite blurry. Some critics rose up against their proposal, contending that some kind of intrinsic property is required in order to consider a process as cognitive, an intrinsic property which only characterizes some internal dynamics of the brain, they argued.

While this first functionalist strategy, based on the so called *parity principle* (see *infra*, I, § 1)[1], got stuck in some metaphysical criticisms about what intrinsically characterizes a process as cognitive, a second wave of thinkers turned the focus on how external resources – especially artefacts – *complement* our cognitive processes.

So, Clark and Chalmers's original argument claimed that our cognitive processes extend into environmental resources, then, there are resources in the environment which *constitute* our mind; by contrast, the new strategy of these second-wave theorists focused on how our minds create synergies with external resources, which enhance and complement our brain-based cognitive capacities.

Even though these two strategies proposed two different perspectives on the relationship between the mind and the environment, they coincided in that the subject of cognition is indeed an *individual agent* (who, then, relies on external resources which *constitute* or at least *complement* her mind). So, more recently, a third wave of arguments questioned both these two individualist approaches to the nature of the mind appealing to the fact that cognition definitely is a *social* fact and so is the mind.

The social turn in the third-wave claims resulted in a determinant point: cognition fundamentally consist in a meaning-making activity. In this work I endorse this view.

So, I argue that cognition in humans mainly consist in meaning production and these production is indeed a social activity. I do endorse, as any extended mind position,

[1] Roman numbers in the internal references included in the text refer to a chapter. Figures and tables, instead, are indicated with Arabic numbers: the first digit refers to the chapter where the figure appears, while the second digit refers to the figure itself (e.g. '6.1' indicates the first figure of the sixth chapter).

an active externalism framework and, for this reason I maintain that the environment plays a central active role in cognition as well. In fact I also claim, in this work, that human communities produce meanings into the ecological space they live within; such meanings, then, crystallize in public representations which are externally deployed in the environment (the cultural niche) the community inhabits; these representations are collectively built and consumed by all the integrants of the group. I shall argue that they both support *collective memories* and *collective beliefs* as emergent mental traits of the community.

Moreover, representations entail normative practices (Menary 2013), which regulates how to use them and, eventually, turns out to be a constitutive element of the mind. In fact, publicly manipulated representations end in the emergence of *canons*, which I shall argue that represent instances of collective beliefs; these canons, ultimately, feed back to the integrants of the community, shaping their minds, their categories and the very meanings they use in their acts of reference. For these reason I shall argue that the very semantics is an emergent property of a community.

Therefore I claim that in *cultural niches*, namely, the culturally shaped environments human communities inhabit, where representations are manipulated in conformity with certain normative practices, there are the conditions for the emergence of a collective mind.

Considering this premise, the thesis will develop accordingly with the following structure.

In the first chapter I shall introduce the debate about the Extended Mind Thesis. It is opened by a short introduction, then it develops in four main sections: the first one deals with Clark and Chalmers's original proposal and takes into account the early criticisms by internalists theorists; the second one focuses on the so called second wave of the extended-mind approaches, presenting the complementarity principle as an alternative to the more classic parity principle as a criterion for the extension of the mind; individual subsections are devoted to a specific focus on the diverse authors who make part of the Second Wave; the third section will be devoted to the analysis of the

Third Wave proposals, with particular attention to the social dimension of the mind; a fourth section will recapitulate the content of this chapter.

The second chapter will be devoted to the concept of *collective mind*: in this chapter I shall advance my thesis about collective mind. After a short introduction, in the first section the difference between the concept of collective mind and some other close concepts will be examined: in fact, I shall focus here, with particular attention, on the specific traits which distinguish the concept of collective mind from the one of collective intelligence and that of self-organization; I shall clarify, then, that there are two main views about the concept of mind which ultimately determine the focus on collective mind: the 'Cartesian' view takes as a term of comparison the features of an individual agent's psychology to detect the same properties instantiated at a collective level; the informational view considers, instead, that the peculiar trait of mind consists in manipulating information, so the authors who endorse this view stress the relevance of public representations as constituents of collective minds. In the second section I shall present my view of how human groups manipulate public representations to produce instances of collective mind; I shall focus, here, on the determinant role that the environment plays in the dynamics of collective cognition. Finally, in the third section, I shall articulate my thesis about collective mind, explaining why the cultural structure of a community constitutes its collective mind; I shall argue that communities produce collective memories and collective beliefs.

The third chapter focuses on the Theory of Niche Construction. This paradigm represented a revolution in the contemporary theory of the evolution of species because it introduced a focus on the feedback loops which emerge between the agents acting in their environment and the environment itself. This chapter focuses on the dynamics proper of the environment which play a crucial role in cognition: after a short introduction to the issue of the chapter, I shall present the general features of the Theory of Niche Construction in the first section; in the second section a shall detail the concept of cultural niche, while I shall devote the third section of this chapter to explain the relevance of the concept of cultural niche for the concept of collective mind I

endorse: in fact, I shall commit here to a scaffolded account of the relationship between the environment and the cognitive processes of the human agents who inhabit it.

The fourth chapter expands the analysis of the cultural niche conceived as a scaffolded mind. It focuses on the use of artefacts into the niche and the construction of an embracing memory all around the community. After a short introduction to the topics analysed along the chapter, the first section focuses on the notion of 'cognitive artefact'; the second section develops some elements of the first one to introduce the concept of 'exogram', an external record of information and a particular kind of cognitive artefact; the third section draws on the concept of exogram to argue for the emergence of a collective memory in a human group, enabled by the use of exograms. Finally, the fourth section introduces the issue of public narratives as the scaffolds which organise normative dynamics into the community; the section argues for a spatial conception of narratives.

The fifth chapter analyses the different kinds of information which a human group manages within its niche. An introduction briefly presents the issue of the chapter, then, the first section makes a focus on the difference between three theoretically close concepts: affordances, constraints and conventions. The section also present the difference existing between the information that we find in an ecological niche, namely, ecological information, and the information that we find in a cultural niche, namely, representations. The second section details the concept of stigmergy as the fundamental dynamics which regulates the feedback loops between the agents and their niche. Particular attention will be devoted to the concept of human stigmergy. The third section focuses on the emergence of public representations into a cultural niche and on the emergence of meanings as well. In the final part of the section I shall stress the meta-cognitive function of representations with respect to the agents who inhabit a certain cultural niche. I shall introduce, here, the concept of mindshaping.

The sixth chapter develops a proposal for a categorial mindshaping as the main meta-cognitive function of public representations. This argument will be supported by the analysis of a concrete case study. After an introduction on the concept of mindshaping, I shall analyse the original notion elaborated by Matteo Mameli (2001); in

the second section I shall present some more recent developments of the notion of mindshaping by Tadeusz Zawidzki (Zawidzki 2008, 2013); in the third section I shall highlight some lacunae of the actual theoretical framework of mindshaping. Finally, in the fourth section I shall introduce the concept of categorial mindshaping and the role of public representations: I shall draw on a case study to defend my point.

Finally, I shall conclude this work summarizing the results I got: the emergent mental properties of a community.

CHAPTER I

Ways of Mind-extending

In the contemporary debate in philosophy of mind a major problem has recently emerged about the nature of cognition, about its structure and its very dynamics: do we think just by means of our brain or, rather, our cognition extends outside of the skull? If so, what would 'cognition' then be? These two questions result eventually in a more radical one: how are we supposed to consider the mind?

In the late 1990s, the philosophers Andy Clark and David Chalmers introduced this question into the debate about the nature of the mind: in a famous disruptive essay (Clark and Chalmers 1998) they suggested that we actually have good reasons to think that the mind *extends* beyond the limits of the skull. They were proposing, in their article, an *active externalism* approach to cognition (Menary 2010, pp. 1-3) and, then, to the very nature of the mind.

Active externalism distinguishes with respect to classical externalism (Putnam 1975) because it is not concerned with semantics; instead, its main concern is the active role the external environment plays in the dynamics of cognition. There are, hence, two ways to consider active externalism: a trivial interpretation focuses on the *influence* that causally active features of the environment have on cognitive processes within the brain; while a robust view of active externalism argues that some features of the external environment are real *constituents* of our cognitive processes.

In a certain sense, the trivial view seems to appeal at most to mere *epistemic actions* (Kirsh and Maglio 1994): this is the manipulation of objects and devices in the surrounding environment, in order to facilitate cognitive processes which take place within the brain. This trivial view of externalism sounds perfectly acceptable to the majority of *internalist* philosophers (those who endorse the thesis of a brain-bounded cognition), because it just affirms that external elements merely *help* cognitive processes, they are not actual *constituents* of cognition. On the other side, Clark and Chalmers defend a constitutive approach to active externalism in cognition: namely, they affirm

that *coupled systems* (e.g. a subject plus a cognitive artefact) are genuinely cognitive entities, and cognition is distributed in the whole system. They argue that the very *coupling link* of the system *is* the cognitive process. This is the point of friction which internalists cannot accept: instead, these philosophers remark that robust active externalism eventually entails a case of *coupling/constitution fallacy* (Adams and Aizawa 2010, p. 67). Around this quarrel developed the so called first wave of the Extended Mind Thesis (Menary 2010a, p. 20), based on the *parity principle*: namely, the principle which considers as equivalent two different phenomena, one of which is located in the brain and the other one is located outside of it, depending on their functional similarity.

Instead, the second wave of the Extended Mind Thesis is based on the *complementarity principle*: this is, internal and external phenomena can be radically different, even though they are complementary for the realization of a cognitive task.

Finally, a third wave of the Extended Mind Thesis has recently arisen, which focuses on the social dimension of cognition and considers the mind as a property of collectives. The aim of this chapter is to explain this development within the Extended Mind Thesis, to present a short introduction to the main positions in this field and to analyse critically the arguments that have been put forward.

1. The first wave

Andy Clark and David Chalmers start considering, in their seminal work, the implications of distributed cognition; so they focus firstly on the role played by epistemic actions in the economy of cognition. They consider a famous (re-elaborated) case from Kirsh and Maglio (1994) about the real and the mental rotation of objects, to make them match in such and such way in this or that socket. They take into account three cognitive situations: in the first one an agent is expected to match geometrical shapes with the correct socket rotating them mentally; in the second case an agent is expected to match the same shapes, which are now displayed on a computer screen, rotating them on the screen until they match the socket; in the third considered case an agent has been equipped with a cyberpunk tool which makes him able to rotate these

geometrical shapes internally displayed by this embedded tool. Hence, both in the first and the third cases the cognitive process takes place within the brain, while in the second case the agent realizes the cognitive task through an epistemic action consisting in the physical rotation of the same geometrical shapes displayed on the screen, so appealing to external resources. This point is crucial: are all these three actions equivalent? If so, considering that both in the first and third cases we clearly are in front of a cognitive process, the second considered case should be treated as a cognitive process as well, even though a part of the task is realized outside of the brain. In words of Clark and Chalmers: «[i]f, as we confront some task, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is (so we claim) part of the cognitive process» (Clark and Chalmers 1998, p. 8).

Therefore, externalizing (at least partially) our cognitive processes entails the *active externalism* briefly introduced in the previous section. Clark and Chalmers argue that the agent and the artefact he is using (in the considered case, a computer screen) jointly form a cognitive system because «[a]ll the components in the system play an active causal role, and they jointly govern behavior in the same sort of way that cognition usually does» (Clark and Chalmers 1998, p. 8). Then, it is remarkable that what makes a part of the world relevant as a constituent of the cognitive process is its *causal role* with respect to behaviour. This is because active externalism acts *hic et nunc*, that is in real time: while the influence of the environment on agents' behaviour in the classical externalism framework is distal, historical, and therefore *passive*[1], in Clark and Chalmers' proposal the relationship between the agent and his environment is proximal and present; and this is why it can support a cognitive process extending in the environment. A typical case is the one of Scrabble tiles rearranged in such or such ways, in order to form this or that word: tiles manipulation would count in this context as part of *thought*, not merely as part of action (Clark and Chalmers 1998, pp. 9-10).

[1] That is, an agent's beliefs concerning his surrounding environment are influenced by his history, namely where and when he has been living, what culture he belongs to, what he has been learning about the world from his environment.

So, the main conclusion of an active externalism approach to cognition is that cognitive processes can be (at least partially) realized somewhere in the environment outside of the skull. But this is not enough to endorse the extension of the very mind into the world, hence Clark and Chalmers go a step further and argue for extended mental states, such as *beliefs*.

Their argument, here, develops through a famous thought experiment of externalization of memory and, then, of dispositional beliefs: the Otto-Inga case. We find here Inga, a perfectly functional agent, who wants to go to see an exhibition at the MoMA museum; she remembers that the museum is on the 53rd Street and goes there. She recalled from somewhere in her memory the information about the museum location, so she has a determinate belief about where the museum is. Then we have Otto: he suffers from a neurological disease which lessens his memory performances. Nevertheless, Otto got a solution for his little problem: he always carries a notebook around with him and takes note on it of any new relevant information. So he relies on the notebook exactly as Inga relies on her biological memory. When Otto finds out about the exhibition at the MoMA he decides to go, so he has a look at his notebook and retrieves the information about MoMA's location, then he gets to the 53rd Street.

Clark and Chalmers argue that, if we accept that Inga has the belief of the MoMA being on the 53rd Street, we should also accept that Otto has this belief too, no matter if in his case the relevant information is stored in an external artefact while Inga retrieves it from her biological memory. This is because, as Inga is confident about the veracity and correctness of the information she remembers, so it is in the case of Otto, whose confident attitude towards the information contained in his notebook makes him automatically accept this information as faithful (Clark 2010, p. 46). Then, if a mental state such as a belief can be stored – so, *extends* – outside the skull, this implies that the mind itself is not brain-bounded but extends in the environment as well.

This argument in favour of the Extended Mind Thesis immediately raised many criticisms, the majority of which attacked the so called coupling/constitution fallacy, pointing at the dubious requirement of *causal relevance* Clark and Chalmers appealed to for the inclusion of external constituents in the cognitive process. More in general, the

majority of criticisms generated from the *parity principle* which Clark and Chalmers endorse in their essay.

1.1. The Mark of the Cognitive and some other 'mental cramps'

So a big difficulty Clark and Chalmers's perspective has to deal with is the one which originates from the parity principle assumed by these two philosophers in their argument: when they say «[...] a part of the world *functions as* [my emphasis]» a process in the head (such as, for instance, Inga's biological memory recall) they are committing to a parity condition which forces them to take into account the necessary structural difference existing between Otto's and Inga's specific cognitive mechanisms. So, they have been attacked on the basis of how coarse-grained or fine-grained their functionalist equivalence is supposed to be. And, if it is to be considered as highly coarse-grained (as it appears to be), many critics (Adams and Aizawa 2001, 2008, 2010; Rupert 2009, 2010b) have pointed at the irrelevance, in that case, of such a similarity.

This is the weak point where many internalist philosophers focused their criticisms. In particular, Frederick Adams and Kenneth Aizawa objected that externalist and transcranialist theories of mind fail in their aim when they defend that cognition (and, then, the mind) extends out of the bounds of the skull (Adams and Aizawa 2001, 2010). They consider that there is no reason to suppose that any cognitive process extends in the artefacts we use in our daily life: after all – they affirm – there is certainly no cognition in the pencil we use to write long computations on a paper sheet (Adams and Aizawa 2010, p. 67). What they argue is that the majority of externalist theorists – and, especially, Clark and Chalmers – fall into a *coupling/constitution fallacy*, namely they erroneously get the conclusion that some cognitive process extends in an artefact just because this artefact is *coupled with* a human agent when this agent is involved in a cognitive task. This criticism points in particular at Clark and Chalmers's definition of an artefact as playing «an active causal role» (Clark and Chalmers 1998, p. 8) in the realization of a determinate cognitive task. For instance, in this internalist view Otto's notebook is just a tool Otto uses to support his memory; but his memory is all

contained within his brain. The notebook is at most a punctual, external stimulus which *helps* memory.

Then, Adams and Aizawa argue that such a generous externalist focus on the nature of cognition and of the mind eventually results in an indefinite set of heterogeneous cases of human/artefacts coupling (Adams and Aizawa 2010, p. 76). So they consider that it would be impossible to study this motley subject through a scientific method, because there are not enough regularities to make possible a real science of brain-tool cognition.

Thirdly, they argue that Clark and Chalmers' position leaves indeterminate the nature of the very cognition: these two philosophers, in their opinion, do not offer any *mark of the cognitive*; this is, any criterion to distinguish between what belongs to the realm of cognition and what does not.

But, what is this 'mark of the cognitive' after all? Adams and Aizawa argue «that cognition is constituted by certain sorts of causal processes that involve nonderived content» (Adams and Aizawa 2010, p. 68); hence, the mark of the cognitive consists in the intrinsic (that is, nonderived) content Adams and Aizawa ascribe to the brain causal processes[2]. Namely, they affirm that while the objects existing in our world (including our cognitive artefacts) get their semantic value just as a second order property we attribute to them, on the other side our brains handle genuine primarily semantic content, a content whose semantic value is intrinsic, so not dependent on external conditions.

Therefore their argument seems to rely on a *semantic* point, namely they focus on the meaning of mental structures whose content is supposed to be intrinsic, while the meaning of the objects of the world is considered as *derived* from this one. So they say that externalists fail when trying to argue for extended cognition because they do not consider the intrinsic semantic nature of a real cognitive process. To this objection an externalist could respond that actually cognition is not just limited to semantics, even

[2] Adams and Aizawa are not so clear about what they mean with the words 'causal process'. They do not seem to appeal to neural structure which should instantiate this or that state of the world. Rather, they seem to refer to Fodor's theory of the mind (Adams and Aizawa 2010, p. 71). In this case, the 'intrinsic content' should refer to concepts as the compositional units of a 'language of thought', a position which is nowadays hardly accepted.

though semantics constitutes an important part of it. An externalist could also say that, as a matter of fact, no one (not even Clark and Chalmers) would refuse considering cognition as *brain-centred* (Di Francesco and Piredda 2012); instead, what externalists suggest is that cognition is simply not *brain-bounded*. That is, none of them would question that the brain plays a main role in cognition, while what they do question is that cognitive processes are only limited to internal states of the brain.

This point is directly related with the first criticism Adams and Aizawa advanced against the Extended Mind Thesis: the coupling/constitution fallacy. They affirm that externalists confuse two different situations: the coupled realization of two essentially separated processes which develop in temporal coordination, with the situation of two different (but coordinated) stages which constitute together a unique process. Adams and Aizawa, then, suggest that externalists are fooled by a misleading perception of two temporally related phenomena erroneously perceived as parts of the same process just because they occur in a linear temporal sequence, like a spectator perceiving a unique and fluent movement in the sequence of the different frames of a film.

Even if these two phenomena were conceived as causally related to one another, Adams and Aizawa argue that there would be no reason to conceive them as constituents of the same process just as my finger switching on a light does not make it part of lighting. In their words: «The neurons leading into a neuromuscular junction are coupled to the muscles they innervate, but the neurons are not a part of the muscles they innervate» (Adams and Aizawa 2010, p. 68). Perhaps this example throws light on the real weak point of Adams and Aizawa's argument of fallacy: it is true that neurons do not make part of the muscles they innervate, but they indeed are constituents of the *movement* of my arm, together with my muscles. In this sense it could be argued, in favour of Clark and Chalmers, that the brain-tool coupling could constitute a cognitive process just as a neurons-muscles coupling constitutes movements.

Finally, the question of the unscientific motley muddle. It could be objected to Adams and Aizawa that their judgement about this point is *tout court* superficial, as Clark does (Clark 2010, p. 50): in fact, they appeal to the heterogeneous framework the works on human-tool cognitive relations are included within. The problem – so I see it –

is that they gratuitously put in a unique heap the ideas of a group of quite heterogeneous authors, such as Donald Norman (1991) and Edwin Hutchins (1995, 2014), Merlin Donald (1997, 2010) and Kim Sterelny (2007)[3] and then they (indirectly) affirm that there is not a homogenous paradigm describing the works of these different scholars, while in neuroscience, for instance, there is a clear paradigm and a determinate framework of reference. But this is only due to the fact that all these scholars focused on diverse aspects of brain-tool cognitive relation – while there is a wide range of different cognitive artefacts (Heersmink 2013) – and all of them did actually find interesting regularities in their respective fields. To make a comparison, Adams and Aizawa's criticism on this point is so ridiculous as if someone were objecting that what neuroscience, psychology and linguistics say about language ontogeny and cognitive relevance constitutes a motley set of incoherent assumptions which is unlikely to meet the requirements of the scientific method.

1.2. Clark's response

In the previous section I presented some of the major criticisms against the Extended Mind Thesis and I also set out some responses in favour of an externalist position.
In this section I shall resume Clark's own responses to internalists' criticisms.

As a first move, Clark questions the very consistency of the mark-of-the-cognitive objection and the alleged coupling/constitution fallacy: Adams and Aizawa are clearly right when arguing that there is no cognition *contained in* the pencil we use for manual computations (see *supra* § 1.1); but we could consider the equivalent case of a V4 neuron

[3] All these theorists focused – each one with respect to his particular interests – on the relationship between mind, body, tools and the surrounding environment. Their respective conclusions and their technical vocabulary are different because different are their theoretical frameworks and their respective aims as well. For instance, as I shall detail in chapter IV, Norman focuses on the concept of cognitive artefact because he is concerned with the relevance of artefacts in computer science; Hutchins treats artefacts as 'external resources' a subject can rely on when dealing with a cognitive challenge: his famous example of the Polynesian sailor orienting her navigation by means of constellation is representative; Donald focuses on the evolutive impact on human cognition of material culture and, especially, of those artefacts and techniques which enabled humans for external storing of their memories (see *infra*, IV, § 2); finally, Sterelny focuses on the influence of the ecological niche some humans inhabit on their cognitive skills: he defended that humans build all around them 'cognitive scaffolds' which assist and harness human cognitive behaviour (see *infra*, III, §§ 2-3).

thinking that there is some peculiar pattern in a stimulus just because it is coupled to a monkey[4]. Indeed, what does not work in this specular situation is the attribution of cognitive abilities to the V4 neuron, as if it could be «some kind of self-contained locus of thinking» (Clark 2010a, p. 83). Certainly cognition is not a 'substance' whose traces can be detected in this or that putative part of a cognitive process; you cannot find it in the pencil but neither in a singular neuron. After all, «[w]hat would it mean for the neuron *or* the pencil to be, as it were, brute factively "cognitive"?» (Clark 2010a, p. 83). This kind of argument makes no sense, as Clark says. This is because, in an active externalism perspective, the status of 'part of a cognitive process' depends on «the kind of role» (Clark 2010a, p. 83) that very part plays within the cognitive system; it is not due to its own intrinsic nature. Also, Adams and Aizawa's criticism on this point is misleading because they discuss against the *cognitive nature* of artefacts allegedly included in the cognitive process; instead, Clark and Chalmers argue for «condition not of "being cognitive" but for *incorporation* [my emphasis] into a cognitive system» (Clark 2010a, p. 84). And Otto's notebook is part of the cognitive process in the sense that it is incorporated into an agent's cognitive activity[5], being crucial for retrieval and use of the information Otto needs.

Then Clark takes a step further against the intrinsic content argument and affords the Martian memory thought experiment (Clark 2010a, p. 89): we could hypothesise the existence of a Martian subject whose perception of the world is organised in bit-mapped images he stores and – occasionally – retrieves from is biological memory; it would be a clear case of intracranial cognition, but it would also be pretty dubious the intrinsic character of such an iconic representation, so similar to Otto's signs stored in the notebook. Would the Martian's memory still count as a cognitive process? It seems reasonable to get an affirmative conclusion. So, what should be the importance of this alleged intrinsic content which characterises genuine cognitive processes? It is not clear at all.

[4] «Question: Why did the V4 neuron think that there was a spiral pattern in the stimulus? Answer: Because it was coupled to the monkey» (Clark 2010a, p. 82).

[5] «[…] the object or process [is] part of the agent's cognitive apparatus» (Clark 2010a, p. 84).

Finally, Clark concludes questioning the very mark of the cognitive chosen by Adams and Aizawa: they focus on neurophysiological and psychological features characterising our human brain; but there is no reason for denying that some other (maybe, alien) organisms could show very different biological routines. Moreover

> «It seems very plausible, for example, that there is no part of the physical universe so devoid of potentially computationally useful properties that that part could not, under some conceivable circumstances, participate as a crucial element in some extended, recognizably computational process, on which some cognitive state of some being supervenes» (Clark 2010a, p. 92).

As Clark remarks, the big problem with Adams and Aizawa's mark-of-the-cognitive approach is that they offer just a 'closed model' of what cognition is supposed to be: if, to spot new cognitive phenomena, you just rely on the causal processes you have so far detected as characteristic features of the (terrestrial) biological cognition, you might not be able to recognise future instances of those new cognitive processes you could discover, because their traits could be radically different from the ones already acknowledged as such (Clark 2010a, p. 92). This is why Clark endorses a much more open detection criterion: «What makes a process cognitive, it seems to me, is that it supports intelligent behavior» (Clark 2010a, p. 92). So, supporting intelligent behaviour is the 'mark of the cognitive' which Clark proposes, a mark which indeed includes the case of Otto's notebook.

What is at stake, now, is then the manner in which something can support intelligent behaviour and Clark himself points at the degree of *complementarity* and *integration* achieved by the components of an alleged cognitive system as a crucial discriminant. It is right from the focus on these two criteria that the second wave of the Extended Mind Thesis arose.

2. The second wave

In the first section of this chapter I mentioned *epistemic actions* as the only extra-body element internalist philosophers could admit within their concept of mind, considering them as external tools which may help brain-bounded cognitive processes. An internalist theorist could easily say that a pencil and a paper sheet are just tools which make computation easier; but, the very computation – he could say – is nothing but what happens in a determinate brain. As the internalist Adams and Aizawa say, the real cognition is characterised by a peculiar feature, namely its (alleged) *intrinsic content*. Here is the point of fracture with Clark's focus: the intrinsic content radically distinguishes genuine cognition from the not genuine one (see *supra*, § 1.1 and § 1.2). I previously explained that this fracture originates from Clark's approach to the Extended Mind Theory in terms of a *parity principle*, that is, he admits a functional equivalence between internal (biological) cognitive processes and external (artificial) cognitive processes (see *supra*, § 1). This is why Adams and Aizawa are so concerned with this version of active externalism: internal and external processes are so much structurally different that it is non-sense to try to consider them as equivalent cognitive phenomena, and such a coarse-grained functionalism approach to cognition is – in their opinion – not worth because of its poor explicative value.

Nevertheless, I do think that internalist criticisms are generally weak – and I said why in § 1.1 and § 1.2 – but a parity principle approach to the Extended Mind Theory does not work so well either, because of its radical functionalism. This is why some other externalist theorists have more recently proposed to approach this subject through a *complementarity principle* and an *integrationist perspective* about cognition.

Among the others, Robert Wilson focuses on an integrationist response to internalist criticisms. He proposes to shift the view and consider *cognitive resources* in cognitive science as *developmental resources* are considered in evolutionary biology (Wilson 2010, p. 174): any creature in an ecological environment certainly has to rely on its genetic legacy to survive, but its adaptation is also conditioned by all the environmental resources it can use to improve its fitness. For instance, the quantity of

food available in a particular habitat might become a considerable advantage for this or that species when building their niche. In the same way, Wilson claims that we should consider as cognitive resources all those facilities that improve our cognitive abilities, which we find available in our environment. This is the *active cognition* argument. So Wilson's point is that our cognition is extended in the sense that we constantly use environmental resources to accomplish the majority of our daily cognitive tasks (Wilson 2010, p. 172). He agrees with Adams and Aizawa that minds firstly are *intentional engines*, this is, they produce and manipulate meanings. But, in contrast with those two internalist theorists, Wilson argues that there is no reason for excluding external resources from being components of such an engine (Wilson 2010, p. 175). After all, the use of external symbolic systems enhances our cognitive capabilities, affording new meaning-making possibilities: just as in the case of Kanzi, the famous bonobo who was claimed to develop linguistic-like capacities thanks to the use of a special keyboard displaying symbols he associated to actions, we can develop new beliefs and desires relying on our linguistic scaffolds, which integrate pre-existing cognitive abilities indeed (Wilson 2010, p. 180). Also, when trying to solve problems in our environment, we often need to recruit external resources, such as the physical manipulation of certain objects: let's think in the common case of a puzzle (Wilson 2010, p. 180). Although, what is crucial in Wilson's view is that

> «[...] it is not simply environmental structures that somehow magically make for cognition, but the causal integration of these with on-board capacities that organisms already have. In general this causal integration can be ontogenetic or phylogenetic, individual or collective, cultural or biological, and the cognitive abilities such integration generates are often genuinely novel» (Wilson 2010, p. 181).

Then, external resources can be enrolled in a cognition process through *causal integration* with an organism's cognitive internal (brain-bounded) skills; this entails the emergence of *novel cognitive capacities*. And all this is possible just because we humans are socially oriented creatures (Wilson 2010, p. 182).

Yet, Adams and Aizawa would clearly object that real intentionality is only proper to brain processes, so Wilson proposes to rethink the very Problem of Intentionality from its roots. He starts considering that in the received view about intentionality (Searle 1995, 2006; Millikan 1984) it is supposed that *mental* representations are characterised by an intrinsic content which makes them mean by themselves: this should be something like the 'essence' of intentionality. In contrast with this perspective, Wilson argues that such an essence does not exist and, then, if we want to understand the real value of representations for cognition, we should shift «our focus from these representational essences to representation in practice» (Wilson 2010, p. 183). Being cognition dynamic in most of its constituents, Wilson points at how we make use of representations in our cognitive routines, so he moves the focus from things such as *representations* to our activities such as *the act of representing* because, «[w]hen cognition is extended, intentionality is extended and the traditional problem of intentionality transformed» (Wilson 2010, p. 184). So, there is no more reason to look for the essence of representation in order to treat minds as intentional engines; rather, we have to focus on the many ways we have to create and manipulate representations, when dealing with the cognitive challenges our habitat affords to us.

2.1. John Sutton: exograms and other elements of cognition

John Sutton is one of the main thinkers who has taken part for the Extended Mind Thesis and he is the one who proposed to analyse the development of this philosophical approach into three different waves, considering the first one as based on the parity principle and opposed to the second one which relies on the complementarity principle. The third wave would then surpass – in his opinion – the very individualist focus in philosophy of mind (Sutton 2010).

Sutton considers that the main criticisms against the Extended Mind just focused on the weakness of the parity principle proposed by Clark and Chalmers since their seminal work, where the Otto-Inga mental experiment had the leading role. As I remarked in the first part of this chapter, dealing with the parity principle implied that

the major criticism was centred in a defence of the peculiar *intentional value* of human (biological) cognition and so the intrinsic content of internal representations. But, as Wilson argued (see *supra*, § 2), in the Extended Mind paradigm we should not focus on the 'traditional' intentionality problem, which tries to explain the very nature of representations; instead, we should consider the act of representing and, then, the representational strategies we undertake as situated agents within our ecological space. So, since we have extended minds, we also have extended intentionality.

Sutton's strategy in defence of Extended Mind is, nevertheless, different: he openly recognises the limits of the first-wave approach based on Clark and Chalmers' parity principle; actually, analysing the Otto-Inga case, he argues that information in biological memory is certainly different from the one Otto records in its notebook. And this is true in a very basic sense: while Inga possesses an *active memory*, Otto does not; pieces of information in Inga's memory can interact and stimulate new behaviours, while this is not the case of the information encapsulated in the notebook, which is in some sense *static*. Namely, the information recollected in an external artefact is active *only when used* by a brain, while biological memory records are always online and they can interact among them even when they are not actually used at a conscious level (Sutton 2010, p. 197). So – Sutton says – Clark and Chalmers' approach is intrinsically weak just because of its functional similarity pretence about cognition; and this is why it is subject to Adams and Aizawa's objections. In contrast with their position, Sutton remarks that the use of external tools to store information for future cognitive needs manifests very different dynamics, if compared with biological memory: following Merlin Donald, Sutton treats Otto's memory in terms of *exograms*, that is, external memory records (Sutton 2010, pp. 189, 197) which have a characteristic appearance and dynamics (they are discrete fashion pieces of information with no intrinsic activity).

But, just because of its peculiar features, external tools play a main role as *complementary elements* in the economy of cognition: they enhance the cognitive capacities of human agents who use them. For instance, exogrammatic information is durable and easily shareable and, in this way, it *enhances human memory* improving its persistence and diffusion. The representative example Sutton chooses to explain the

effectiveness of external memory is the one of actors' performances in English baroque theatre. In the famous Globe Theatre, for instance, actors used to represent various characters in different plays taking place in a short period; they have to remember a lot of different texts and situations and all these data would have been impossible to manage if they had to rely just on their biological memory. So they developed a spatially distributed mnemonic system: they disseminated cues within the stage which were useful to remember entire lines with just a quick look at a script with general information about what they had to do in a particular moment; also, they could rely on hearing cues in the structure of the performance (Sutton 2010, pp. 202-203). In this way they had the possibility to successfully handle a huge quantity of information which would have been unmanageable for someone dependent on his own biological memory.

So, we scaffold our environment with external structures which help us to handle all the information we need and we cannot afford with our poor biological capacities. A similar case which supports this approach to Extended Mind as a scaffolding technique we implement in the space we inhabit, is represented by nautical maps or stars reference system as a complementary tool for navigation (Hutchins 1995, 2014): we can rely on the relative position of our boat with respect to the Polar Star to know which way the North is.

Finally, Sutton builds on Clark's account of cognitive technologies and argues that there is no reason to think of them as *necessarily external* structures: we could treat as a cognitive technology any internalised system of symbols, such as language (Sutton 2010, p. 207). In this sense, he sees as a good example of internalised symbolic scaffold the typical 'palaces of memory' so characteristic of the Renaissance: the scholars of this period had to memorise a lot of information they could not carry around with them; to make this job easier, they developed a very sophisticated art of memory which implemented different strategies to make memorisation more affordable (Eco 2013). As already common among Latin rhetors, these scholars used to 'build' imaginal spaces such as palaces, squares, streets which were called *loci* (that is, places). To each *locus* was associated a determinate information; then, the structural relations among the different *loci* of this virtual map were used as cues for memory by the scholar who was 'moving'

through this imaginal space. It is evident the similarity of these structures with the Globe Theatre system of spatial cues as a support for biological memory. In some sense, giving a visual structure (such as a city architecture) to the scaffold a scholar wanted to memorise, is just a particular way to re-design the environment, changing the cognitive task the agent has to accomplish: it is a strategy to make a certain problem fit better in our cognitive schemes, like an intelligent use of space (Kirsh 1995). So, re-designing our memory in a determinate scaffold, changes the very structure of our cognitive routines and is a complement for our biological cognition dynamics.

Using external structures (even when internalized) as a support and, then, a *complement* for biological memory is just a first argument in favour of an Extended Mind paradigm, stronger than the one imagined by Clark and Chalmers at the beginnings of this philosophical movement. In the next section I shall present the *cognitive integration* approach proposed by Richard Menary as a second argument in favour of this second wave.

2.2. Richard Menary: the cognitive integration

Richard Menary is another main philosopher supporting the Extended Mind Thesis, but within the framework of the second wave. Just as Sutton, he sets out some criticisms to the first-wave approach: he explicitly points at the problems related to the parity principle and the coarse-grained functionalist focus proposed by Clark and Chalmers (see *supra*, § 2.1). He endorses a more enactive approach to cognition and focuses on how our bodily activities, neural processes and external tools interact in our cognitive practices. So he argues that cognition has to be considered as a *hybrid practice* which features coordination between internal and external processes, when a subject is carrying out a cognitive task (Menary 2010a, p. 228).

Menary endorses, then, the Theory of Manipulation (Menary 2010a, § 4), that is, human cognition is possible when we manipulate our environment to achieve some cognitive task, so we integrate both internal and external resources (which are explicitly recognised as structurally different). He says that there is no good reason to think of

cognition as an a priori brain-bounded activity. On one side, it seems pretty gratuitous to establish an essentialist mark of the cognitive as Adams and Aizawa do: although neurons are intuitively considered as intrinsically cognitive, it is nevertheless clear that their activity is due to the migration of potassium and sodium ions along the nerves, elements to which no one would in principle reduce cognition (Menary 2010a, p. 230); it seems then more reasonable to abandon the speculation about an alleged mark of the cognitive, to focus rather on cognitive processes as those dynamics which allow the accomplishment of a cognitive task. On the other side, if someone would argue for the irrelevance of external resources in the economy of cognition, she should explain why we humans spend so much time in so costly activities, such as the production of external representational systems. Menary's answer is that we use all these environmental resources as a decisive support for our biological cognitive capacities: for instance, we always write out our mathematical problems to work on them and find an appropriate solution, we do not solve them in mind due to the huge amount of data we should manage; this speaks in support of cognitive integration, because «[i]f cognition is bounded by the brain, why do we not complete all these cognitive tasks, and many others like them, "in the head"?» (Menary 2010a, p. 231). In this way Menary suggests that external resources are *necessary constituents* for human cognition to be effective.

There are two main approaches to cognitive integration: one which relies on a phenomenological account of our embodied cognition (Gallagher 2005); then a second approach which considers cognition as distributed (or extended) in the environment (Hutchins 1995; Clark and Chalmers 1998; Sutton 2006a). The first one relies preferentially on bio-causal dynamics between an agent and his environment, while the second one prefers to consider how an agent's cognition is supported by cultural scaffolds and specific tools. In this context, there are various kinds of bodily manipulations we can undertake in the surrounding environment, such as biological coupling (extended phenotypes and sensory motor contingencies); epistemic actions (see *supra*, § 1); regulative actions (some behavioural uses of language and norms) and cognitive actions (the use of external representational systems according to determinate normative

practices). All these are cases where the manipulation of the environment is crucial to accomplish a certain task (Menary 2010a, p. 237).

Manipulations are possible because we develop *cognitive norms* (Menary 2010a, p. 238), namely, systems of instructions to use our environmental resources when integrating cognition: Otto can appeal to the information-bearing signs collected within his notebook just because this cognitive dynamics takes place in a wider cultural and normative system. So, what Menary's account adds to the extended-mind style arguments, typical of the first wave, is an explanation of how it is possible for human agents to use so many diverse representational systems to integrate their cognitive routines. This is possible because «[any] manipulation of an external representation is normative» (Menary 2010a, p. 238); this means that any practice we learn is a *method* for manipulation of representations that we acquire, because cognitive norms are the instructions which guide a certain cognitive practice. Moreover, this practice has to be conceived as normative in so far as we use the method to learn how to *correctly accomplish* it. All this is crucial for Menary's definition of the cognitive:

> «The pen and paper or the CPU, keyboard, and monitor are not themselves cognitive (nor are the sodium ions traversing nerve fibers in the brain), but the creation and manipulation of the external vehicles and the coordination of internal and external vehicles is» (Menary 2010a, p. 240).

So, what actually makes external resources cognitive is the very integration of them into our cognitive routines through the correct application of the cognitive norms we developed, that is the method.

2.3. Other externalists: Michael Wheeler and Mark Rowlands

Michael Wheeler inserts into the debate about the Extended Mind Thesis with a proposal in defence of *extended functionalism*. Although his contribution is chronologically closer to the second wave of arguments in favour of the extended mind,

he should be considered as a representative of the first wave. This is because he appeals to the parity principle instead of arguing in favour of complementarity. Nevertheless, even though his proposal conceptually fits better in the first-wave paradigm, I am analysing his work here because he deals with some of the second-wave theoretical pillars such as the complementarity principle and the like.

He defends the importance of a parity principle approach because he considers that it constitutes the only way to successfully support an extended cognition focus with genuine cognitive basis. The reason for this move is that Wheeler agrees (surprisingly) with Adams and Aizawa about the weakness of the *causal coupling* requirement as a criterion for mind extension (Wheeler 2010, p. 246): his point is that in no way the mere causal intervention of a tool into the development of a cognitive process makes such a tool a constituent of that process. And this is – in his opinion – the very weak point of any embodied/embedded approach to the extension of the mind, due to the fact that he considers that these kinds of perspectives conceive tools as mere objects manipulated by the cognitive system composed of brain and body (Wheeler 2010, p. 247): an embodied/embedded theorist could have, in principle, no problem at all in considering tools as a simple 'help' for cognitive activities, which are genuinely displayed just within the body scheme (Gallagher and Zahavi 2008).

But, albeit embodied/embedded theorists do not consider tools *per se* as constituents of cognition, they include them within the body scheme as co-opted elements in support of certain kinds of cognitive dynamics (Rizzolatti and Sinigaglia 2006; Gallese and Sinigaglia 2010).

Then, Wheeler argues that if we want to effectively defend a genuine mind extension outside the skull/skin limits we have to be committed to an extended functionalism perspective: namely, relying on Clark and Chalmers' parity principle, it is possible to argue for the real extension of cognitive processes outside the body; this is because Wheeler's argument relies on *multiple realizability* of mental properties as a consequence of functionalism (Wheeler 2010, p. 251). In this sense, any system implying different components which regardless of its structure work in all the relevant cognitive respects as the brain or the body does, is to be considered as functionally equivalent to

them: for instance, in his functionalist approach, what crucially makes a systemic extended state work as a mental state is all the set of causal relations supporting systemic sensorial inputs and behavioural outputs, like perception and the correspondent reaction (Wheeler 2010, p. 249). This is quite different with respect to the mere causal coupling requirement, because extended functionalism does not appeal to the occasional use of tools or other props as instances of extended cognition; it rather focuses on *systemic dynamics* of recollection and use of information and their functional equivalence when they are considered as a whole. In sum, his point is that the parity principle does not refer to the functional parity between the intrinsic dynamics of a concrete external tool and the intrinsic dynamics of a certain internal cognitive process; rather, the whole system has to be considered as a functional peer or not with respect to a general definition of the cognitive in functionalist terms: the functionalist equivalence has to be established between an instance of extended system and an independent standard (Wheeler 2010, p. 253). So, functional equivalence is not to be thought of as a relationship between an external tool and a concrete internal dynamics like, let's say, spatial memory; instead, real functional parity has to be sought in the comparison between a whole cognitive process in, let's say, World-1, which is completely internally deployed, and a whole cognitive process in World-2, which is partially carried out through the involvement of external resources. If functional parity is acknowledged between both systems, this means that cognition can involve even external elements as genuine constituents, not as mere instances of causal coupling.

So applying an extended functionalism approach is, in Wheeler's opinion, the unique way to defend a genuine extended cognition framework. This would be the unique way to argue for external elements as *genuine constituents* of the mind, intrinsically cognitive when coupled with the subject, while a complementarity approach would not be equally effective.

Nevertheless, Wheeler is wrong when considering that any embodied/embedded approach to mind extension fails advocating it just because its reasons allegedly rely on causal coupling arguments: he erroneously thinks that all the embodied/embedded theorists agree to treat external tools as *objects* for a *subject*'s cognition, instead of

considering them as real constituents of those cognitive dynamics. Rather, if we examine Menary's integrationist approach, we have to acknowledge that it entails the inclusion of external coupled elements as genuine components of cognitive processes. And this is, actually, the case of writing out mathematical problems to solve them through the physical manipulation of written symbols; an emblematic case of embodied/embedded theoretical failure for Wheeler (Wheeler 2010, p. 247), a good example of cognitive integration for Menary (Menary 2010a, p. 240)[6].

We can go deeper into the analysis of this problem considering Mark Rowlands' approach to the *extension of consciousness*. He argues that conscious experience does not supervene on intracranial processes; rather it rises when in perception we undertake a *disclosing activity*, whose vehicles are generally external to the brain (Rowlands 2010, p. 285). To illustrate this concept, he proposes to take into account a classical example of Maurice Merleau-Ponty (1945/2003): when a blind person uses his cane to get situated in the surrounding space, the cane itself stops being a mere *object* and becomes a *vehicle* of perception. In his words:

> «when the stick becomes a familiar tool, the world of the tactile objects withdraws, it does not begins anymore at the epidermis of the hand, but at the bottom of the stick. We are tempted to say that by means of the sensations produced by the pressure of the stick on the hand, the blind builds the stick and all its different positions; then, these positions mediate an object to the second power, the external object. […] Both the pressures on the hand and the stick are not data anymore, the stick is not an object that the blind would perceive anymore; instead, it is a tool 'through' which he perceives. It is an

[6] This is because Wheeler considers that the use of written symbols as a help for calculations is not to be considered as an isolated element which may be compared with internally deployed mental calculations; the right term of the equivalence should instead be the whole system [brain + mathematical notations + pen + paper] considered *with respect to the cognitive task*, not in comparison to some determinate internally-located process. In this sense, Otto's notebook is not to be considered functionally equivalent to Inga's biological memory as such; rather, the system [Otto + notebook] should be considered functionally equivalent to Inga's cognitive routines with respect to a third term: the cognitive task of remembering a certain address.
For this reason Wheeler rejects the embodied/embedded approach as an unsatisfactory response to internalist claims.

appendix of the body, an extension of the corporal synthesis» (Merleau-Ponty 1945/2003, pp. 216-217)[7].

So Rowlands defends that in such cases the external tool used by an agent becomes, as described in Merleau-Ponty's example, a vehicle of a *genuine perceptual experience* (Rowlands 2010, p. 287), because it is decisive for the very production of such an experience. And this is a quite diffuse phenomenon at many levels for any cognizer: let's think for instance of saccadic eye movements or sensorimotor activity, where the mechanical movement of a certain part of the body plays a determinant role in the production of a particular phenomenal experience (Rowlands 2010, pp. 288-290). Also, there are well known empirical proofs supporting the thesis that the use of a tool to achieve a concrete goal changes the very structure of our body scheme as it is conceived by our brain (Iriki, Tanaka, and Iwamura 1996; Rizzolatti and Sinigaglia 2006).

It is clear that in all these examples each tool used by an agent is a *real constituent* of his conscious activity, it is not a mere external 'help' for some brain-bounded processes, rather it is a main part of a cognitive integration case. This is why Wheeler's view about the need of a parity-principle approach to defend the constitutive role of external elements of cognition is wrong: embodied/embedded views do not rely on a causal coupling argument; instead, they argue for a *structural change* in our spatial perception operated by the use of those tools and, then, for a *constitutive role* of such objects in our conscious dynamics.

2.4. Robert Rupert: some more arguments against the Extended Mind

We have seen so far the main arguments proposed by those externalist thinkers who have contributed to the so called second wave of the Extended Mind Thesis. It is now useful to consider some more criticisms to this approach, elaborated by Robert Rupert.

His view is mainly focused on two points: the first one is concerned with an account of what a cognitive system eventually is and what reasons we have to consider

[7] This English translation is mine.

this or that entity as a genuine constituent of it; the second one is concerned with an analysis of the relationship between language and cognition, whose aim is to criticise the language-based argument in favour of the Extended Mind Thesis (that is, considering language as an external component of cognition).

Rupert has a position in many respects close to Adams and Aizawa's worries: first, he considers that any possible system composed by a human subject and a set of external linguistics resources, which make it extend into the environment, cannot support any of those skills which cognitive science could be interested in, because of its limited longevity or integrity (e.g. it is a scarcely durable unit whose constituents are not integrated enough); second, he argues that classical arguments in favour of the extended mind generally rely on a *dependence-reasoning*, namely the inference that from the dependence of thought on a concrete factor deduces that such a factor is a *constituent* of the thinker's cognitive system (Rupert 2010, p. 326). This last point appears to be a finer version of the coupling/constitution fallacy we have seen in § 1.

So Rupert argues that relying merely on a coupling criterion is a too liberal condition for the existence of an extended cognitive system, after all – he says – coupling situations are pervasive in our daily experience and in many cases just trivially influent on a determinate subject's cognition. Also, this does not entail at all that the coupled elements are actual components of the alleged system: for instance, when a child experiences the linguistic features of her environment it is true that these elements *drive* in some ways her behaviour, but she has no active control on them[8] (Rupert 2010, p. 327); the relationship is unidirectional, then they do not form a real system.

Rupert's point, here, is that for two elements to be components of the same system means that they are actually *interdependent* in a strong way: I could say that the reason why I am able to see is that the sun sheds light on the things I have all around me, but this does not imply that the sun is a component of my visual system (Rupert 2010, p. 328). Nevertheless, here Rupert seems to be fooled by the assumption that every extended approach to cognition has to follow Clark and Chalmers' argument on

[8] This last judgment being quite problematic, in my opinion, considering that any speaker has a certain control on the linguistic elements in his environment, given that linguistic elements often consist in communicative exchanges.

constitution. But the system taken into account by those two authors is an instance of a very peculiar kind: namely, they are worried with the possibility of an *external active memory*, while the sun-eye system for visual perception does not represent an instance of extended cognition in any sense: in fact, the sun (and therefore sunlight) is for the eye just an external *precondition* for visual perception, it certainly is not a *component* of visual perception itself. A similar situation characterises Rupert's analysis of Wilson's view: he definitely does not understand (at least, so it appears!) why Wilson decisively defends the idea that individuals are cohesive entities, concrete spatio-temporally bounded organisms and at the same time argues in favour of extended cognitive systems (Rupert 2010, p. 329). Here Rupert, just as Adams and Aizawa, is missing the central point of the different approach of the second wave perspective on extended mind with respect to the first wave: the parity principle is changed for the complementarity principle. Wilson does not endorse any functional equivalence between external and internal processes; instead, he proposes an integrationist view about how our internal capabilities are *complemented* by external structures. So, there is no contradiction there!

Perhaps for Rupert even complementarity principle entails too loose requirements for a process to be legitimately considered cognitive. His position, after all, is that cognition requires *interdependence* of processes and an alleged external cognitive process should satisfy such interdependence criterion. Nonetheless, I think that considering interdependence as the key factor for determining what is cognitive and what is not would ultimately be a too demanding criterion: in fact, it would certainly exclude from the range of the cognitive the majority of the integrationist cases proposed, among the others, by Menary. However, it is a common experience for many people the impossibility of realizing long and complex calculations without relying on external supports such as pen and paper and the like; the introduction of algebraic notations afforded novel mathematical possibilities for human minds; the use of symbolic systems in general allows cognitive functions otherwise inaccessible. The point is that all these cases, even if none of them requires interdependence but just complementary coordination, they are nevertheless cognitively relevant: in fact, employing such external tools decisively *redesigns the cognitive task* (Kirsh 1995). This is, in my opinion, a

sufficient condition for the inclusion of external resources into the range of cognitive dynamics, at least in the integrationist paradigm.

Then, for what concerns the short durability and the wide variety of many (putative) cognitive systems, Rupert sees an insurmountable problem in the absence of a clear *criterion of demarcation* of an extended cognitive system. This is an evident re-elaboration of Adams and Aizawa's mark-of-the-cognitive argument I criticised in § 1.1. So, he concludes that an alleged cognitive science of extended systems would be devoid of a concrete individual to analyse in a coherent study (Rupert 2010, p. 332).

Finally, the language-based argument.

Many externalists (Dennett 1991; Clark 1997, 1998, 2003; Sutton 2006b, 2006a; Sutton et al. 2010; R. A. Wilson 2010) have proposed that language actually constitutes an extension of our cognitive activity and it is the ultimate cognitive artefact (Rupert 2010, p. 325). It is our most complex apparatus for mind-extension. The reason is that it is seen by externalists as a way we have to off-load information into content-laden external symbolic systems (Rupert 2010, p. 335): for instance, every time we write down a list we are creating an external storage for some information we have in mind. The idea is that such an external system can provide a subject with new cognitive capacities and novel schemes to manage information (Rupert 2010, p. 338): for instance, new logic rules or conceptual structures. So here the question becomes, for Rupert, whether the use of an external linguistic apparatus adds something more to internal cognitive processes or not. More precisely, the main problem is whether linguistic external structures have some information of their own or, rather, they are *merely symbolic tools* whose content is *derived* from internal cognitive dynamics (Rupert 2010, pp. 339-340). Rupert argues that there is no reason to believe in an autonomous content of external linguistic structures (Rupert 2010, p. 336). It is quite clear that here we have a slightly more subtle version of Adams and Aizawa's argument of the *intrinsic content* (see *supra*, § 1.1): external linguistic scaffolds can just afford us logic structures to manipulate a semantic content which has been previously produced by internal processes of representation. He just admits, following Carruthers and Boucher (Rupert 2010, p. 340), that external linguistic scaffolds provide us with a *model* for internal processes.

Finally, Rupert accepts that external symbolic systems such as language, even though they do not feature non-derived semantic content, are anyway a useful *support* for cognition:

> «The internalist approach claims that external language provides the subject with augmented computing power and new cognitive strategies by offering to the subject fixed mental units that serve as stand-ins for her own thoughts, units which can then be the object of further reflection and manipulation» (Rupert 2010, p. 344).

So, he concludes that such an external scaffold is nothing more than a useful 'help' for cognitive processes, which he places (only!) within the brain. It is worth noticing, here, that Rupert endorses a view which conceives the possibility for our brain to use external tools which serve its internal cognitive dynamics, for instance, making some computation easier. What he is not willing to accept is the conclusion that this relation supports eventually an extension of the considered cognitive system. Instead, as many other internalist theorists, he is open to consider the role of external elements as tools for a more smooth implementation of brain-bounded cognitive processes.

Then, here it is relevant to notice that Rupert's position should be reconcilable, at least in principle, with an integrationist perspective on mind extension, such as Menary's position or Wilson's view. Actually, all of Rupert's criticisms seem to effectively attack only externalist approaches based on functional parity. It is singular to see that because of nearly the same reasons (namely, complementarity) which pushed second-wave theorists to argue for mind extension, Rupert is disposed to refuse any real possibility of genuine mind extension. The integrationist view might eventually reconcile these different perspectives.

3. The third wave

The third wave of the Extended Mind Thesis features a quite novel aspect if compared with both the first and the second waves: it shows a clear switch from the individualist

focus to a *group perspective* on cognitive dynamics. In fact, although the two previous waves have a different view of cognition and cognitive systems (namely, oriented to the parity-functionalist approach in one case and the complementarity-integrationist approach in the other one), they both converge on the assumption that the subjects of cognition are *individuals* (Cash 2013, p. 61), whose capacities can be artificially *enhanced* (and so extended) by the use of external tools or *integrated* by the synergy between internal processes and environmental resources. This is evident in Clark and Chalmers' famous example of Otto's mnemonic capacities improved by the use of the notebook, where both he and the neurotypical subject Inga are examined as individuals. And still Wilson manifestly proposes an individual organism as the subject of cognition, whose cognitive skills can be integrated by environmental resources. Certainly, all the internalist thinkers defend an individualist position about the subject of cognition as well.

So, the radical change of focus introduced into the debate by the third wave theorists consists in considering the mind as a result of collective practices emerging in groups, obliterating the thorny question of the *bounds of cognition* (Cash 2013, p. 62): there is no individual subject of cognition anymore, so cognitive processes span into the community. This entails two hypotheses: a) individual minds are collectively moulded; b) groups develop minds of their own. So, for what concerns the first hypothesis, the attention is centred on those *normative practices* which emerge from many individuals acting within their group, then developing conceptual structures which eventually mould individual minds; we find a representative example in the emergence and development of linguistic rules (Cash 2013, p. 64) or in the legal system (Gallagher 2013, p. 6). Instead, the second hypothesis focuses on mental features of the very groups, such as collective agency, collective personhood or collective beliefs (Huebner 2014, 2013; List and Pettit 2011; Theiner 2011; Szanto 2014).

Let's analyse the first hypothesis. Shaun Gallagher, introducing his enactivist focus in the quarrel about the extended mind, argues that individuals' cognitive dynamics are strongly influenced and partially constituted by certain social structures which characterise our public environment. Institutions, social norms and cultural

practices are an actual cognitive scaffold we have all around us: each time we act within the law, we are appealing to external conceptual schemes which do not merely exist in individual minds; they rather exist as shared structures, as *mental institutions* (Gallagher 2013, p. 6).

Gallagher attacks the received individualist view arguing that it offers a 'Cartesian' image of the mind because, even though in such perspective the mind *extends* into the surrounding world, it is nevertheless conceived in terms of those processes taking place within the brain (Gallagher 2013, p. 5). Functionalism implies that, if a process X carries out the same function carried out by process Y, they have to be considered equivalent under that respect; then, if Y is recognised as a cognitive process because it carries out a certain function and X carries out that function as well, X should also be considered a cognitive process. In this case, the requirements (the functions) to be a mind are established with respect to the good old brain-bounded cognitive processes. Since this is the term of comparison in the received extended view, we find that all the extended cognitive processes analysed so far by externalist theorists are conceived as an artificial reproduction (by means of external tools) of some psychological features, such as memory, priming effects, dispositional beliefs and desires, or the phenomenology of emotions. Instead, Gallagher observes that these few criteria do not exhaust all the spectrum of cognition and appeals to enactivist dynamics: he argues that there is no reason to exclude a priori an external structure from our cognitive processes just for being not so reliable or automatically endorsed as those processes characteristic of our biological routine, disagreeing here with Clark and Chalmers's functionalist criteria (Clark and Chalmers 1998, p. 17). Such external structures do play a main cognitive role implementing the solution of many problems we face in the daily life. For instance, contracts represent external records of memory about an agreed-upon decision; they have a mental nature (because they are constituted of beliefs, desires and more propositional attitudes), but they do not exist within individuals. They exist in a public space where they are shared and publicly recognised by the community:

> «Contracts are institutions that embody conceptual schemas that, in turn, contribute to and shape our cognitive processes. They are not only the product of certain cognitive exercises, but are also used as *tools* [my emphasis] to accomplish certain aims, to reinforce certain behaviors, and to solve certain problems. Institutions of property, contract, rights, and law not only guide our thinking about social arrangements, for example, or about what we can and cannot do, but *allow us to think* [my emphasis] in ways that were not possible without such institutions» (Gallagher 2013, p. 6).

It is crucial he says that contracts (and mental institutions in general) *shape* our cognitive processes. This means that normative practices *per se* feedback on individuals and, in this sense, mould their behaviour and make part of their socially extended cognitive processes. They are, after all, an instance of a very complex *epistemic action* (Gallagher 2013, p. 8): for example, when a judge delivers his ruling he appeals to an entire *system of laws* existing in the community and uses it as an epistemic tool to elaborate his decision; in the same way, a scientist relies on some *scientific practices* (accepted by the scientific community) when doing research. The point here is that both the judge's decisions or the scientist's statements *would never be possible*, *nor even conceivable*, without such institutions.

So intelligence should be considered, in his opinion, as a problem-solving and behaviour control mechanism, rather than a mere set of propositional attitudes and psychological states; it should be conceived as a cognitive engagement with the environmental resources we share and then cognition does not just reduce to a movement from inside the subject, but from the outside towards the subject as well (Gallagher 2013, p. 6).

Yet, Gallagher affirms that these mental institutions exist because there is a number of people which are actually *cognitively engaged* with such public information and recognise its value. This point seems to echo John Searle's collective intentionality (Searle 1995, 2006) and, then, presupposes the ultimate problem of how to explain the emergence of this collective feature.

Eventually Gallagher defines cognition as «any interaction or engagement that produces meaning for the agent» (Gallagher 2013, p. 8) and meaning production is never

an individual enterprise, but a participatory one, where sense making is collectively shaped by super-individual norms, like those mental institutions I mentioned above. Since such institutions enable many of our cognitive processes, they are actual constituents of cognition just as the shovel I use is (jointly with my muscles) a constituent of digging (Gallagher 2013, p. 10). The point is that in so far as mental institutions are conceptual schemes that we collectively share in our communities, they work exactly like those tools which, when used, extend an agent's body scheme (see *supra*, the end of § 2.3).

In a recent article, Menary criticised some of Gallagher's conclusions arguing that it is not proper to define the so called mental institutions as *constituents* of mind (Menary 2013). He suggests, instead, that such institutions are external tools and scaffolds which *integrate* our mental activity: he says, for instance, that Gallagher's representative model of the legal system is a clear example of an external scaffold of information a lawyer can appeal to when working on a case, being this a manifest instance of integration. He sees no reasons for defending a constitutive approach to such external scaffolds; he declares, in fact, that they are evolutionary strategies which improve an agent's problem solving skills and, even though they are active components of a cognitive process, minds do not extend within them.

So the main disagreement between Menary's and Gallagher's perspectives concerns how we should conceive mental institutions and what is their role in the economy of cognition, whether they are *constituents* of human minds or rather they are mere *integrations* of human cognitive skills. This entails a subtle but deep divide between these two perspectives: in effect, both Menary and Gallagher agree on the point that human cognition relies in some way on external resources; they both agree on a description of intelligence as a problem-solving activity as well. Nonetheless, on one hand, Menary focuses on individual agents whose cognitive performances depend on *individual integrations* of external resources (for instance, the use of algebraic notations affords *me*, inasmuch as I am an individual subject, novel possibilities of calculation otherwise impossible); on the other hand, Gallagher stresses that problem-solving also implies a cognitive engagement with the available environmental resources and, in his

opinion, cognition certainly is also a matter of *meaning-making*, which entails a collective endeavour in the constitution of the meaning supporting the existence of mental institutions. For instance a judge's ruling is not merely integrated by the appeal to the whole legal system; rather, *it is constituted by* the action of appealing to a certain system of laws, which eventually is meaningful just because of a collective attribution of meaning.

Actually, the real disagreement on a constitutive claim about institutions, arises between those theorists who endorse an *enactive* view of cognition and those who argue for an *embedded* focus about cognitive processes: in fact, Gallagher considers mental institutions as components of cognition in their own right because he eventually has a *semantic* focus on cognition, which implies that a certain action gets its proper sense only in its proper context. Nevertheless this does not mean that some concrete prescription gets its own meaning (and then becomes possible) just because the correspondent legal system makes sense of it in terms of definitions (e.g. "There is a fine for parking here" needs a legal system defining the terms 'fine' and 'parking'); instead, that determinate prescription is to be conceived as an epistemic tool which affords *new conceptual possibilities* to the agent who uses it. To employ an example of Menary's (Menary 2013, § 2.1) against him, having algebraic operations available within our mathematic system did not only offer us a tool for more accurate calculations, merely integrating our previous individual skills; instead, it *changed* our very minds providing us with *new concepts* we did not have before. Then, when Bryce Huebner argues that social structures like institutions do not constitute our minds but they just complement them, he is criticising Gallagher's enactivism in favour of a socially *embedded* interpretation of our cognitive routines (Huebner 2013, p. 15). Just like Menary, he does not see – in my opinion – Gallagher's central point: institutions are not mere tools which particular agents can use at a certain moment to accomplish a concrete task (for instance, delivering a juridical decision); rather the collective itself is producing and using such institutions to enable determinate conceptual possibilities. So, while in Gallagher's perspective the real subject of cognition is the very collective, this is not the case in Menary's or in Huebner's proposals.

Moreover, we have a sub-question here: the disagreement about the ultimate nature of the cognizer (namely, whether we have to consider collectives as such or better individual agents) is tightly connected to the vision these theorists have of cultural practices and, then, social structures like institutions. The point is that many embedded-view partisans seem to conceive cultural practices as kinds of representations that some agents are projecting in a public space; instead, as Edwin Hutchins[9] proposes (Hutchins 2011), they are dynamic elements independent of any concrete brain:

> «Cultural practices include particular ways of seeing (or hearing, or feeling, or smelling, or tasting) the world. Cultural practices are not cultural models traditionally construed as disembodied mental representations of knowledge. Rather they are fully embodied skills. Cultural practices organize the action in situated action. Cultural practices are emergent products of dynamic distributed networks of constraints. Some constraints may be internal and mental (some of these are perhaps consciously experienced, but most are implicit and affectively charged), some constraints arise from the mechanics and physiology of the body, some constraints may be provided by engagement with material artifacts and some from interactions with social others» (Hutchins 2011, p. 441).

Then, cultural practices constitute a dynamic structure which is emergent and independent with respect to any individual. In fact, «few of the dynamic loops that link people to their environments are invented by the people who exploit them» (Hutchins 2011, p. 441). They eventually are systems of *public constraints* which *shape* the behaviour and the cognitive strategies of individual agents who have access to them. This is coherent with Gallagher's definition of cognition as the very production of meaning, which is definitely a collective enterprise (Gallagher 2013, p. 8).

So collective production of meaning is a key point in the third wave of the Extended Mind Thesis, a point echoing the famous Vygotsky's studies about the ontogenesis of concepts as a product of collective communicative engagement (Krueger

[9] It is true indeed that Hutchins defends a *culturally embedded* view of cognition (Hutchins 2011). Yet, he argues that cultural structures such as constellations do not just integrate but are *constituents* of such a peculiar cognitive activity like navigation (Hutchins 1995, 2014).

2013; Meini 2012). A semantic structure, once emerged, has to be conceived as those constraints Hutchins refers to (see *supra*, his last large quotation). This perspective opens new possibilities about mind-extension, projecting us to the concept of *collective mind*: in effect, we have no reason to hypothesise that an extended (nor a collective) mind should reproduce those Cartesian features we are so used to take into account when analysing individual minds (Gallagher 2013, p. 5). So, building on this conception, I would suggest to approach the collective mind affaire considering it as an emergent system of semantic constraints which then organises individuals' behaviour[10].

Nevertheless, the second hypothesis about collective cognition I introduced at the beginning of this section focuses on whether collectives develop minds of their own or not in terms of a more standard model of mind featuring its typical Cartesian attributes, namely whether they show agency, personhood, beliefs, decision-making operations and all those psychological features characterising individual minds or not.

Because collectives are generally analysed in terms of *juridical personhood*, it is easy to find in the literature examples where collective agents are firms, institutions or associations (List and Pettit 2011; Huebner 2014). Then, there is a certain generalised openness with respect to all the *functional* attributes of the mind that can be acknowledged to group-agents. For instance, a company can be considered as having its own aims apart from its integrants: the JuicyFruits Co.[11] might quite likely have the aim to control the possible reaction of fruit-collectors unions, in order to avoid any strike against a planned salary reduction; the JuicyFruits Co. might even undertake some concrete actions to pursue its objectives and it will surely be considered accountable in a court for such actions, being it a juridical person; nevertheless, the majority of its integrants might easily have no particular aims (maybe not even beliefs) with respect to the unions and the stance to adopt.

Instead it is quite problematic the point of the alleged *consciousness* characterising a collective: many theorists are fine with a description of group consciousness in terms of *awareness* (e.g. having agency or beliefs), but almost no one would attribute to a

[10] I shall clarify this point in the next chapter.
[11] Obviously, a fiction company.

collective any kind of *phenomenal consciousness* (List 2015, pp. 2-3, 8). It seems pretty fine, in this context, to overlook the question of a phenomenal consciousness in collective agents because it sounds definitely metaphorical to talk about such a thing like 'the headache of the city' or 'the pain of the country'. This is because these experiences (at least when considered in literal terms) necessarily appeal to *physical sensations* which are clearly impossible separately from a nervous system, and indeed in a collective agent the only nervous systems (in a literal sense) are those existing within the bodies of its integrants. There is no reason to look for the alleged functional equivalent of a headache in a group-agent and it is even less clear why we should take into account a metaphorical use of some terms like 'headache' in this context. More in general, it is quite problematic applying a Cartesian conception of the mind to collectives without ending trapped in a fractal dilemma[12]. Instead I suggest that we should look for the novel and peculiar features of group minds, since they likely display a characteristic nature which might be very different with respect to the model of an individual mind.

4. A backward glance

Looking backwards at this theoretical panorama, it is worth to see how the debate about the alleged mind extension started focusing, at first, on the very nature of cognition, arguing for some intrinsic feature or against it. So, in the first wave the discussion focused on *where* cognitive processes happen, whether they can extend in tools we use in our daily life or rather they are brain-bounded. For that all the debate drew attention to the validity of functional parity criteria.

The second wave switched, then, the discussion to the *integration* of our internal cognitive dynamics with some external resources we find in our environment. This entailed overlooking the question of an alleged intrinsic feature of genuine cognition, to devote instead attention to those external elements which do not substitute but *complement* our cognition.

[12] The attribution to the whole of the same features of its parts.

The third wave moved a step further from an externalised and integrated cognition towards a more liberalised view of the relationship between individual brains and the surrounding environment, overturning the focus and arguing for a collective dimension of cognition. This point is crucial to highlight how, from the original extended mind claim, a new debate arose about the possibility of a *collective mind*.

In the next chapter I shall focus on some main points of the contemporary debate about collective mind and, after that, I shall make my own proposal of what a collective mind is. I shall remark the novel emergent features which characterise a collective mind and make it structurally different from and individual mind. Indeed, I shall discard the first-wave strategy based in the parity principle because, as I highlighted all along this chapter, it results to be ineffective for a strong defence of mind extension. Instead, in the next chapter I shall keep grounding my proposal on two main conceptual pillars: a) the *complementarity of external resources* for human cognition, especially for what concerns the manipulation of artefacts and the dependence on the surrounding environment; b) the *meaning-making criterion* as a distinguished feature of cognition, which marks the difference between a mere instance of collective intelligence and the more interesting case of a collective mind.

CHAPTER II

What a collective mind is

In the last chapter I have been reconstructing the theoretical panorama concerning the issue of the extended mind: from the first classical formulation of the problem which, relying on the so-called *parity principle* (Clark and Chalmers 1998), defends that cognitive processes may extend into the artefacts we use to accomplish a cognitive task, to the third-wave arguments, which overturn the question and focus on the relationship between a certain subject and her environment, eventually advocating cognition as the outcome of collective dynamics.

I discarded the first-wave approach because of its slippery metaphysics of the mind which, however suggestive, is nevertheless ineffective against the most subtle internalist criticisms (Rupert 2004, 2009, 2010a, 2010b). Then, I do commit to the second-wave most characteristic requirement for mind extension, namely *complementarity* between a certain agent and the artefacts/tools she uses to carry out cognitive tasks, but I insert myself into the third-wave theoretical framework, considering that I also commit to Gallagher's definition of *cognition as a meaning-making activity*, therefore, a collectively-grounded activity. In fact, as the third wave of the extended mind theory reminds us, we humans are a social species and our cognitive abilities ultimately develop within a highly socialized environment. Also, our main evolutive linchpin is represented by our extraordinary capability in developing new technologies to improve our fitness in the continuously changing environment we inhabit. Technologies are of various kinds: they are *artefacts* (see *infra*, IV), tools which work as artificial extensions for movement, objects manipulation, perception enhancement and cognition in general, boosting our natural capacities; they are also theoretical frameworks (Gallagher 2013, p. 6) which permit us to activate different focuses on the surrounding environment, systems of information which help us in handling our life in that environment. So, we humans usually create groups and develop cooperative strategies to deal together with common problems. Our cognitive activities

always rely on this collective background and this is why I contend that it is reasonable to develop this premise into a collective account of the mind.

Therefore I shall formulate, in this chapter, my own proposal about the collective mind which emerges in a human group. My central thesis is that the very cultural structure emergent in a human group can legitimately be considered as its emergent mind, a *collective* mind. As I claimed in the last chapter (see *supra*, I, § 3), drawing on Gallagher's definition of cognition as a primarily meaning-making activity, I endorse a concept of collective mind as an emergent system of semantic constraints, which feeds back to the integrants of the community it has emerged from and, then, it organises the behaviour of those individuals.

To defend my position I shall firstly distinguish the notion of collective mind from other close concepts, like those of collective intelligence and self-organised system. Fundamentally, for my defence of a collective mind I shall argue that: a) cognition implies a *meaning-making* activity shared by all the integrants of a determined community; b) this collective meaning-making activity is enabled by the collective use of a *shared environment* the considered agents inhabit; c) this shared environment features a plenty of resources the agents can use to complement their cognitive dynamics, such as *artefacts*, tools and other scaffolds; d) among all these elements, collectively shared *representations* are crucial because in public representations meaning crystallizes; e) representations entail *normativity*, which regulates how to use them and therefore turns out to be a constitutive element of the mind; f) in *cultural niches*, namely the human environments where public representations are manipulated in conformity with certain normative practices, there are the conditions for the emergence of a collective mind.

Hence, in this chapter I shall sketch the theoretical schema in which all these elements combine to explain the emergence of a collective mind in a human group, while I shall devote the following chapters to clarify in detail the concepts that appear all along this second chapter. In the next section I shall remark the difference between the concept of collective mind and some other close notions: I shall highlight that a *semantic level* is what ultimately marks the difference between many natural collective

phenomena which entail some degree of intelligence and, on the other hand, the notion of collective mind; then, I shall focus on the issue of representations. Eventually, in the third section, I shall deal with the issue of *culture* as an instance of collective mind and with the emergent properties of a collective mind.

1. From collective intelligence to collective mind

1.1. Collective intelligence

In cognitive science research the concept of 'collective' has been getting more and more attention in the last few years. Evidence of the crucial role of joint actions in the emergence of new cognitive dynamics is increasing in many fields (Malone and Bernstein 2015). Instances of collective intelligence have been detected in biology, economics, social and cultural practices. The main argument supporting the collective intelligence hypothesis (that is, the hypothesis that groups, inasmuch as they are collective entities, can manifest phenomena which may be labelled as intelligent) leans on the relationship between the different individuals integrating a group and the physical space they act within when accomplishing some tasks (Gordon 2015, p. 44): the point is that any integrant of a group tries to keep up with her particular objectives, gradually changing at any new step its local environment and, then, indirectly influencing the actions of the other integrants of the group. Local changes affect individual behaviour and cause effects at the macroscopic level of the group behaviour. There is no need for an explicit shared goal, there is no need for agents' awareness of the global structure they make part of. Also, collective intelligence does not even *necessarily* refer to a *cooperative* link among a certain set of agents (Malone and Bernstein 2015, p. 3), rather it refers to macroscopic effects of group activity in a determinate space in general, which include cooperative links as well. Therefore, albeit cooperation is not a necessary criterion for the occurrence of collective intelligence[1], cooperation-based intelligent systems are the most frequent in

[1] In fact, competitor (but interdependent) agents can produce an intelligent solution to a common problem as an outcome of their individual actions.

nature: for instance we have an evident case of cooperation-based collective intelligence in fish schools and bird flocks, whose collective movements are due to simple local rules about distance and direction among individuals any agent follows[2]. In this context, it is defined as 'intelligent' any behaviour which improves the organization of the considered group and gives to it any kind of adaptive advantage with respect to an hypothetical less organised 'previous stage': for example collective movements in fish schools give each integrant more possibilities to survive a predator attack with respect to solitary fishes, while bird flocks make flying easier for birds which make part of such an organization reducing air friction in the flight. So in this context it is considered as intelligent any kind of successful behavioural adaptation to the environment and, then, it is defined as collectively intelligent any kind of successful behavioural adaptation attributable to the group.

A key consequence of this definition is that the subject of the environmental adaptation is now the collective, not its integrants anymore. In this sense, 'adaptation' might simply refer to an improved capacity for shock-absorption the collective entity develops: it is the case of the stock market, when it is considered as an instance of collective intelligence (Lo 2015). The stock market – under these conditions – might be seen as an interesting example of collective intelligence because it is clear that its components are not linked through cooperation, rather they are *competitors*; so the 'aims' of the group do not coincide with the aims of its integrants, because adaptation into a stock market implies (at least in theory!) the best goods allocation and prices management, with no mercy for the majority of particular interests. Of course, it is arguable if we have here a genuine instance of collective intelligence: Andrew Lo, for instance, defends that the 'intelligence' of the market consists in detecting the responsibility of a certain company in a determinate case (e.g. a public damage or a bad investment) which can affect its economic value; his point is that the market as a whole 'smells' the fault and punishes it, devaluating the responsible economic actor (Lo 2015,

[2] Cooperation, here, is not to be understood in the strict sense of the intentionalist focus: fishes do not share beliefs, representations, objectives and the like (see *infra*, V, § 1). Rather, cooperation here indicates that the integrant of a certain group are not competitors; instead, they evolved social-life routines which enable them to deal better with the attacks of a predator.

pp. 23-25). But it is not clear for me in which sense we could talk here of collective intelligence; rather, in my opinion this case seems to show a trivial case of avalanche-effect: *some investors believe that* a certain company they have invested in at some point got wrong in its actions and then it is going to lose value, so they start selling and this provokes an avalanche-effect. They can be right or they can be wrong (Lo 2015, p. 25ff); in any case I think that it is worth noticing that the decision is made by *individual investors* who concretely estimate risks, the only subjects to whom we paradigmatically apply a psychological framework in a literal way. In this context, then, it sounds barely metaphoric to say that the stock market believes anything.

In the analysis of this example it is possible to appreciate a common difficulty for the theorists of collective intelligence: not any system whose macroscopic features depend on its integrants' local activities shows, just for this reason, an intelligent behaviour. The case of the stock market is emblematic: who is here the ultimate cognizer? Does the stock market really form any kind of belief about a possible state of the world? It is maybe more likely that the effects we can appreciate at the macroscopic level of the stock market behaviour are merely due to *the sum of the particular beliefs* instantiated by the integrants of the system, who eventually result to be the only genuine cognizers. Moreover, the unique feature completely attributable to the market as a system is its tension to *adaptive organization*, but there is no reason for us to treat every instance of adaptive organization as a cognitive activity *stricto sensu* (see also Huebner 2014, p. 221). Actually, adaptive organization is what any ecosystem as a whole tends to, and it sounds kind of weird to consider an ecosystem as a cognitive agent, unless one is willing to endorse something like James Lovelock's Gaia Hypothesis (Lovelock 1979).

So, the inference which heads to such a conclusion about the emergence[3] of a collective intelligence, is likely due to the theoretical influence of an easy analogy with

[3] 'Emergence', here, is to be understood as the *logical* relationship between those properties that we can appreciate focusing on a collective as a whole and, on the other hand, those properties visible at the individual level of analysis: for instance, a certain order is an emergent property of a collective, a property that we can notice when considering the group level, while that order might be absent when focusing on individual routines. I do not commit here to any particular position about the possible *ontological* relations between individual and collective properties.

collective organization of social species in nature (Gordon 2015). The most famous example of this kind appeals to social insects colonies as intelligent organizations. Especially, Edward O. Wilson and Bert Hölldobler (E. O. Wilson 1975; Hölldobler and Wilson 2009) developed the concept of *superorganism* to label ant colonies conceived as the real selective unit in Darwinian terms, for what concerns the environmental adaptation in these insects. An ant colony is considered an instance of collective intelligence because the colony as a whole is able to manage pretty complex problem-solving situations which no ant could manage acting solely. An ant colony is used as a representative example of emergence with a cognitive output because its macroscopic behaviour emerges from simple algorithms individual ants follow locally. These algorithms are nothing more than *phylogenetic adaptations* which made the colony behaviour smoother and more efficient in its environment. An example of emergent behaviour, for what concerns ants, could be mapping foraging areas (Bonabeau et al. 1997): what generally happens in an ant colony is that many foragers get out the nest seeking for food sources in different directions; when one of them finds something interesting, it gets back to the nest signalling all the trail dropping off pheromones; each signalled trail starts attracting more and more ants because of the pheromones; each new ant following the track reinforces the chemical signals to the food source; when a determinate food source starts decreasing, ants stop reinforcing the trail. At a macroscopic level it *appears* that the colony is moving its 'tentacles' as an octopus looking for food, but what we can appreciate at the collective level is nothing more then the result of a series of many local avalanche-effects originated by just few individuals in a stochastic process. Although this last consideration makes the ant colony *prima facie* quite similar, in its interior dynamics, to what we have remarked with respect to the stock market case (that is, emergent macroscopic phenomena from local avalanche-effects), nevertheless we have here an important element to consider: because an ant colony pursues, as a unit[4], the collective fundamental objective of survival (which certainly implies many problem-solving activities such as seeking food sources, keeping

[4] The members of a certain colony share the same genetic inheritance and the same chemical identifiers (each colony has a particular 'smell') and they share the same communication system as well.

itself safe from predators attacks and many other similar examples) into a competitive environment, it is clearly involved in a cognitive activity; instead it is not clear at all that a stock market implements any kind of cognitive activity, because it seems to be more an *ecological space* (see *infra*, III, § 1) where different actors compete for survival than a unified system which yields an instance of collective intelligence. So those who argue in favour of a description of the stock market as an instance of collective intelligence are maybe erroneously assimilating a common case of bottom-up self-organization of a system to a form of collective intelligence (see also Huebner 2014, p. 222).

This pushes me to deal now with two different questions: a) why self-organization is not enough to yield collective intelligence; b) what cognition eventually is in this framework.

1.1.1. Self-organization and collective intelligence: not synonyms at all

The label 'bottom-up' refers to the movement of organization which characterises the emerging order of a system from the local activity of its elementary components. Self-organization is then a feature which refers to a determinate *state of the system* at a certain time T_1, with respect to a previous stage at time T_0. It refers to a state of things more than to a set of real-time reactions a dynamic system smoothly implements to respond to its environment mutations. This is why self-organization characterises not just complex living systems, but any autocatalytic system, even chemical reactions in many kinds of molecules (Kauffman 2000). However we daily experience a lot of self-organization phenomena even in our human social environment. Let's think of the common case of pedestrian traffic direction on city sidewalks: there are no explicit top-down rules which organise the collective movement along a crowded sidewalk, even though people start organising their movement, adapting their speed to avoid collisions with other pedestrians and we can appreciate emerging 'streams of traffic'. Does this entail that is there something like an alleged collective intelligence of pedestrian traffic which organises its integrants? Certainly, the most obvious answer seems to be a categorical

"No!". This is because we have here just a clear case of self-organization whose emerging order is the mere outcome of many local responses which develop towards a certain 'equilibrium', like in the example of the stock market; with no need for any form of central control, but with no common objectives either. This emerging order is just a state of things, an epiphenomenal consequence of basic local reactions: no feedback loop develops here. Moreover, walking pedestrians do not constitute a durable group, so there is no reason to postulate a collective cognitive organization of an alleged 'group' of pedestrians. This last point marks a clear difference with the case of an ant colony: in fact, an ant colony is a well-defined group, both for genetic inheritance and communicative system; also, an ant colony does develop a durable feedback loop among that determinate group of agents and the environment they inhabit. This is certainly not the case of a crowd walking along a transitory space like a sidewalk.

1.1.2. What cognition is in this framework

Cognition, instead, is generally conceived as an *activity*, a problem-solving activity (Malone and Bernstein 2015, p. 3): we could recognise cognition in a certain subject managing its life and continuously bumping into obstacles (that is, the problem) and looking for the best solution to keep its life up (namely, the solving stage). We have, then, an intelligent activity any time a subject is involved in a problem-solving situation; and we have collective intelligence any time the problem-solving activity is attributable to a determinate group. Moreover, there is a remarkable difference between basic cognition and high-order cognition (see *infra*, V, § 1): basic cognition consists in the problem-solving activity common to any species of intelligence, it consists in an appropriate reaction to environmental stimuli and it does not entail any meaning-making activity; instead, high-order cognition consists in problem-solving activity characterised by meaning-making (Gallagher 2013, p. 10) and, therefore, by the use of representations (like Gallagher's examples, see *supra*, I, § 3). I shall argue that this kind of high-order cognition (namely, meaning-making activity) is what ultimately justifies the transition from mere collective intelligence to collective mind.

Therefore, in this theoretical framework, the concept of 'collective mind' entails a further step into the debate, which now tends to drive the attention to the thesis that groups can show *mental features* that are not reducible to their members. Then, the concept of 'mind', if compared with the concept of 'intelligence', seems to be charged with much more theoretical implications because, for some authors, it appeals to a *psychological reality* (or, as I shall defend, a *semantic reality*) which is not necessarily implied by the concept of 'intelligence'[5]. For instance, in the literature about the collective mind, the research focuses on features like collective beliefs, collective intentions, the agency of the group or the responsibility of collective entities, as well as the authorship of collective works produced by a determinate group. So, I will use 'collective intelligence' to refer to something that collectives *do*, while reserving 'collective mind' to characterize certain features that collectives may *have*. This is because the former definition refers to what a collective does to improve its fitness, what *strategies* it implements, while the latter refers to a set of *structural properties* defining the alleged cognitive activity of a determinate collective.

1.2. Collective mind

The contemporary debate about collective minds, as I shall outline throughout this section, is mainly focused on two different approaches to the structural properties which distinguish a minded entity from a mere case of intelligence: the first one defends a more 'Cartesian' view of the mind, namely more centred on classical psychological features which generally characterize individuals (List and Pettit 2011; List 2015); the second one argues for an informational definition of collective minds, that is, a functional perspective which highlights the importance of the structures of information manipulated within a group by its integrants (Theiner, Allen, and Goldstone 2010; Theiner 2011; Szanto 2014). Let's see them in detail.

[5] Flock behaviour is an intelligent performance of collective adaptation, but we have no reason to presuppose *collective beliefs* attributable to the very flock considered as a whole and not reducible to its integrants.

1.2.1. The 'Cartesian' view

The problem of collective mentality originates in 'Cartesian' terms (that is, in comparison with individual psychology features) because we commonly attribute psychological features to many collective nouns in our speech: it is quite frequent to hear phrases like "What does the Party think about that point?" or "What are the plans of this football team for the next season?" or "What will the Company do to increase its profit?" or "Our Journal is delighted to publish your manuscript". So, what is at stake is whether all these examples are mere ways of speaking which synthetically refer to the various members of a group or rather they betray the real existence of autonomous psychological features which are proper of collectives and, then, it would make sense to speak about collective minds in a literal sense.

This is a crucial point. We humans are social beings and we are very often involved in social dynamics which seem to appeal to some kind of 'common aim'. Also, our routines commonly acquire the structure of collective actions, such as group-hunting or environment engineering, building houses and villages or tools for common use within the group, as well as the implementation of defence strategies. All these examples seem to entail a certain kind of 'intentions sharing' which eventually allows the emergence of coordination in the group. This is particularly evident in those peculiar social structures characterizing our social practices which are called *institutions*. Institutions are all those things which so densely populate our social sphere, such as assemblies, judges and tribunals, governors and other political or religious figures, companies, political parties and unions. And, ultimately, the very law system, the net of conventions which regulate how these 'social entities' work. The philosopher John Searle defined all this totality as the 'social reality' (Searle 1995), postulating a cognitive mechanism proper of human brains and responsible for what he called 'collective intentionality'. His idea is that every time a group of people can share an objective, that group is actually sharing intentions, that is, semantic and meta-semantic information about their action as a group, which produces a *social ontology* (that is, all those things like 'money', 'borders', 'marriages' etc. which just exist because of a previous convention valid for the

integrants of a certain group) (Searle 2006) and implements a *collective psychology*, namely collective beliefs, collective desires and collective projects to realize them.

Nevertheless, Searle does not explain in any naturalistically acceptable way how groups produce collective intentionality; after all, this is not his pretension. In his account, collective intentionality is just a *postulate* to make sense of human ability for gathering to reach common objectives. Even though, because of its psychological implications, this postulate appeals to the ultimate question of whether there is a collective mind or not in a group of people, responsible for collective intentions[6]. Christian List and Philip Pettit, for instance, argue in favour of *collective intentional states* (List and Pettit 2011, *passim*) such as beliefs and desires which are proper of collectives and which are not reducible to the individual intentional states of the components of a determinate group. Their argument takes into account the difference between individual ruling of a certain judge and the ruling made by a tribunal composed by a group of judges, this last considered as an instance of collective judgment or collective belief. List and Pettit claim that it is quite common in collective decision to incur in a 'doctrinal paradox', that is to get a different decision depending on whether it is premises-based or conclusion-based (List and Pettit 2011, p. 44ff); this would show the difference between a 'majority decision' and 'the group decision': while the majority of the members of the tribunal, when separately considering the premises, would make a certain ruling, the court's verdict might instead be different, offering then an instance of collective belief (Huebner 2014, p. 142). Also Margaret Gilbert prefers to defend

[6] As Christian List and Philip Pettit remind us (2011, p. 11), the theoretical history of collective agents likely dates back to the Middle Age, when a juridical debate based on Roman law developed about the concept of *legal person*, with the aim of justifying the rights and duties of collective entities such as monasteries and cities. That debate was mainly fed by Cicero's idea of the State as the totality of *leges* and *mores* which constitute the *ius*, namely the abstract and transcendent structure of the State conceived as the fundamental law – the Constitution, in modern words – of a certain group of people. In modern philosophy, this debate developed in the well-known theorization about the origin of the State and the nature of its power. Famously Thomas Hobbes in the 17th century argued in favour of a contractual constitution of the State as a power emerging from a primitive agreement of a group of people, and Jean-Jacques Rousseau theorized the General Will as the will of all the members of the political community operating for the common good. This metaphoric account of the State as a collective entity, animated by a general will emerging in a political community, characterized most of the political thought of the 19th and early 20th century (List and Pettit 2011, p. 8), but the contemporary discussion about collective mind is looking for a more rigorous account of the psychological features of groups (if they have any).

collective mentality in terms of the attitudes of a 'plural subject' (for example, a union) whose integrants «*jointly commit to act as a single body* to bring about some end» (Huebner 2014, p. 141). So Gilbert seems to claim for a preliminary agreement the components of a group reach about a certain common aim (for instance, better work conditions), and this common aim would represent their shared intentions and it would also be a mark of the collective mentality attributable to the group.

These examples are willing to explain the nature of a collective mind in terms of collective intentions (beliefs, desires, etc.) attributable to an alleged 'plural subject' and, for this reason, they use as a term of comparison individual psychological features, considered as the model of the mental.

1.2.2. The informational view

Another main theorist of collective mind like Bryce Huebner tries instead to justify this claim appealing to the capacity of a collective subject to produce and to consume determinate *representations* (Huebner 2014, p. 151) in a way that is not reducible to its individual integrants: for instance, a team organized in different sub-systems, each one analysing different elements of the environment and generating the relative information, would be in Huebner's opinion an instance of collective mind if and only if there is not a central *decision maker* which recollects, processes and uses all the information gathered by the diverse sub-systems which compose the collective agent (Huebner 2014, p. 157). For Huebner, then, a genuine collective agent must meet the fundamental requirement of a distributed recollection, storage and use of representations. In this sense he individuates two representative examples for what he defines as a *minimal collective mind* (Huebner 2014, § 9.4) and for what he labels *maximal collective mind* (Huebner 2014, § 9.7). The first case refers to a well-known example of collective cognitive system, that is, a honeybee colony (Seeley, Kirk Visscher, and Passino 2006; Celli 2008; Passino 2010; Trianni et al. 2011): in a honeybee colony different foragers get out seeking food; when any of them finds a food source, it comes back to the hive and attempts to attract the attention of the other bees starting its waggle-dance; a waggle-dance is an *iconic*

representation (a map) of the food source position with respect to the hive and the sun; individual foragers produce these iconic representation which are publicly shared in the hive and are consumed by all the colony (Huebner 2014, p. 231). The second most representative case is acknowledged by Huebner in the structure and cognitive dynamics of the European Organization for Nuclear Research (known as CERN) and the scientists who integrate it. Within the CERN all the different sub-systems are brought up-to-date about what is happening in the rest of the structure through internal reports, memos, emails (Huebner 2014, p. 251); also, the information is locally collected and evaluated, then the rest of the people are informed by means of internal short reports; these reports work like representations and, being publicly shared, they last for years as *external representations* and they are used as a 'guide' for the investigation along the years (Huebner 2014, p. 254). These ones are actual public representations which are collectively produced and consumed by the whole CERN community. A third midway case of collective mind is represented by groups featuring *transactive memory* (Huebner 2014, § 9.5): this is, for instance, when a married couple who have spent many years together can reciprocally complement their personal memories; or we find another example in a group of co-workers where each of them stores in her memory some pieces of information and all them are aware of *who knows what*. Any member of the considered group could store some piece of information and, through meta-memories acting as 'directories' (Huebner 2014, p. 234), she would be able to appeal to that member of the group who stores this or that piece of information. In this way, all the information available for the group results to be a collective possession and memory manifests itself as a collective mental property.

Finally, other authors like Georg Theiner, Colin Allen and Robert Goldstone (2010), propose a computational theory of cognition as well, arguing in favour of group minds through a strategy which takes its move in the extended-mind framework: they consider that accepting the extended-mind approach we accept at the same time a blurry distinction between the very cognitive role of all the components of an extended cognitive system, for instance between the diverse constituents of a brain-body-artefact system; considering the intrinsic social nature of human beings, there is no reason for

excluding the other people from the cognitive system each of us constitutes when incorporating tools as cognitive artefacts (Theiner, Allen, and Goldstone 2010, p. 380). But they pose a fundamental condition for group mentality: cognition within the group cannot simply be an «unstructured aggregation of individual cognition» (Theiner, Allen, and Goldstone 2010, p. 382); instead it has to be the result of *a clear division of labour* among the members of the group, which implies *cognitive specialisation*. This is why they appeal to the case of transactive memory as well, considering it as a representative instance of collective mentality into a group agent. Also, they consider that many cognitive capacities like attention, problem-solving and (more importantly) *memory*, are information-based processes and, then, they are independent from a concrete support for their implementation. They apply a functionalist perspective to conclude that information can be manipulated and transmitted beyond the neuronal level (Theiner, Allen, and Goldstone 2010, p. 383). On this line of arguments, these authors keep claiming that we do not attribute at the neural level of a subject the same properties we attribute to this subject when considering her in terms of a personal agent (neurons do not have beliefs!) (Theiner, Allen, and Goldstone 2010, p. 384). By analogy, it seems gratuitous to attribute at the group level the same mental properties we experience at a personal-agent level: there is no reason, in principle, to expect to find at a group-agent level something like a 'consciousness' similar to that one we see at an individual-agent level of analysis. By contrast, collective beliefs represent a different issue: human groups, I claim, indeed features collective beliefs, but they are not to be conceived in 'Cartesian' terms; they needs instead new criteria which, in my opinion, are of the same kind of those of Theiner and colleagues. So, these authors argue in favour of the emergence of distinctive and novel mental features when talking of group-agents and collective minds. They argue that for a group to be considered a cognitive system in its own right three conditions must be satisfied: first, the system has to be integrated and to produce functional gains with respect to its integrants acting separately; second, the group produces systemic effects working as a single causal whole; third, the functional gain is possible when the system enhances the individual capacities of its constituents or, rather, it manifests novel emerging skills (Theiner, Allen, and Goldstone 2010, p. 390).

Then, considering these two conceptions of the structural properties a collective mind features (that is, the 'Cartesian' conception and the informational conception), the ultimate question possibly is: what does really matter when looking for a collective mind? While the 'Cartesian' view is interested in psychological similarity with the individual-subject model, the informational view focuses instead on the very structures of information which a certain group features, no matter what kind of support implements those patterns, because the only important thing is the information itself and the functional gains it produces for the group. For this reason, those who commit to the informational view do not look for any kind of individual internal states likeness when seeking the mental properties of groups. Rather, they focus on how determinate patterns of information are built and manipulated within a human group (e.g. structures of shared representations).

Certainly, I commit to the informational view and I shall advance my own proposal in the next section.

2. Public representations, meaning-making and the collective mind

As I said before, I claim that 'collective intelligence' refers to something that collectives *do*, such as the adaptive strategies they undertake and the concrete effects that these yield; while 'collective mind' refers instead to something that collectives *have*, such as some structural properties which ultimately produce mental features like attention or memory. I have also claimed (see *supra*, II, § 1.1), following the contemporary literature, that collective intelligence is an emergent feature of groups whose integrants are cognitive agents operating in a determinate *physical space* and interacting among them mainly through *local actions* that entail *indirect effects*. So we have collective intelligence in flocks or in social insects because each individual, though acting locally, modifies the common environment affording *new information* to its fellows, influencing their behaviour, eventually yielding emergent macro-phenomena. We humans indeed produce instances of collective intelligence when we associate in groups. Moreover, depending on the specific structural properties characterizing human collectives, many authors argue

that we humans are capable of producing instances of collective minded entities as well: as I have explained, some of them underpin the argument for a collective mind on the comparison with individual psychological features; some others, instead, individuate the mental properties of a collective in the informational structures the group features. I commit to the latter view: the informational perspective I have outlined above (see *supra*, II, § 1.2.2).

I already committed to Gallagher's definition of cognition as a problem solving activity which implies meaning-making (Gallagher 2013, p. 10), at least inasmuch as we are dealing with high-order cognition (see *supra*, II, § 1.1.2). So, I argue that we find a genuine collective mind when a human group is involved in a meaning-making activity through the manipulation of some shared structures of information: that is, representations.

In fact, as Huebner argues, the most fundamental commitment of the informational conception of the collective mind is that something like a *collective structure of information* has to be produced and used by all the members of the group to consider that very group as a genuine *minded* entity, for instance, publicly shared representations. This is because representations are conceived as external public vehicles for some particular piece of information which is not individually possessed by anyone. Hence, both in the case of the honeybee hive minimal mind and in the case of the CERN maximal mind, we find that representations are publicly shared by means of artefacts: the waggle-dance in the former case; memos, emails and internal reports in the latter case.

The information which structures the mind, in this informational view, is recollected by the integrants of the group during problem-solving activities, namely, cognitive activities; but high-order cognition entails meaning-making; so ultimately the information carried by representations is the very meaning produced by the agents involved in cognitive tasks; therefore, representations have to be conceived as *deposits of meaning*. But, I would remark here that meaning production is never an individual enterprise; it is rather a collective one. It is a collective enterprise in the same sense of collective cognition, which gradually emerges from local actions operated by individual

agents (see *supra*, II, § 1.1.2). I claim that such collective production of meaning is supported by the collective manipulation of public representations by the agents which constitute a collective mind.

Coherently with these theoretical premises, I shall propose here that the collective structure of information which distinguishes a human group as a collective mind case is its very *culture*, which I define as the totality of public representations collectively produced and consumed by all the members of a community. That is the structure of information which gives coherence to the group and guides its actions.

Consequently, I argue that a human group organizes information all around it in a structured *scaffold* of public representations, which are collectively produced through the gradual contribution of each member of the community and which are publicly shared and consumed by all of them; these public representations set up *social constraints* in form of normative practices and self-regulation norms for individual cognition, coordinating it at a collective level with all the components of the group. This makes such representations the linchpin of a collective mind. Moreover, the collective production of these representations implies a functional gain for the integrants of the group, because collective representations are public tools which boost the cognitive capacity of each individual (especially – I shall argue – in *categorization*).

Considering that publicly shared representations are the vehicles which enable the collective production and manipulation of meaning among the integrants of a collective mind, it ensues that the very semantics is collectively established and shaped through the manipulation of these representations, and eventually it constitutes a novel emerging property of collectives.

In the next section I shall explain how publicly shared representations emerge as deposits of meaning by means of local manipulations, operated by the integrants of the considered group into the physical environment they inhabit.

2.1. The emergence and manipulation of public representations

2.1.1. Multi-agent dynamics in an ecological space: the concept of stigmergy

I previously said, in this chapter, § 1.1, that the emergence of collective intelligence is strictly connected with the relationship existing between the individual agents which integrate a certain group and the physical space they act within. The point here is that this physical space is not an 'abstract space', but a real *ecological space*. This is crucial, because, as David Kirsh remarked in a famous paper (1996), in an ecological space each creature (each agent) tries to preserve itself and to improve its fitness within the habitat and, in order to do that, such an agent can adapt to the environment, migrate to a new habitat or modify the environment itself, adapting it to its own needs. We have to consider that, when trying to adapt the environment to itself, that is, building its own ecological *niche* (see *infra*, III), an agent undertakes a lot of actions as a reaction to environmental stimuli, each of which results in *a change of the structure* of that very habitat.

This implies that the habitat is a continuously changing space where any agent modifies (as a side effect of its particular aims) the environmental stimuli available for all the other agents sharing the same space. This is the indirect relationship existing between two or more agents operating in the same habitat. A good example of this is a beaver building its dam: it realizes a series of actions to improve its fitness in its habitat (that is, building the dam); these actions change the structure of the environment, modifying the *selective pressures* it affords to the other agents acting in the same ecological space; the dam represents a new *constraint* into the environment which limits the range of possible actions for some other agents living in the same habitat of the beaver, for instance, some fishes in the river where the beaver lives, while it affords new possibilities for other agents of the system such as water-birds which can now fish in the pond. This is important because what we have here is just a beaver attempting to redesign its habitat to make it *more comfortable* for it, with no intention to affect the life of other possible agents sharing the same space.

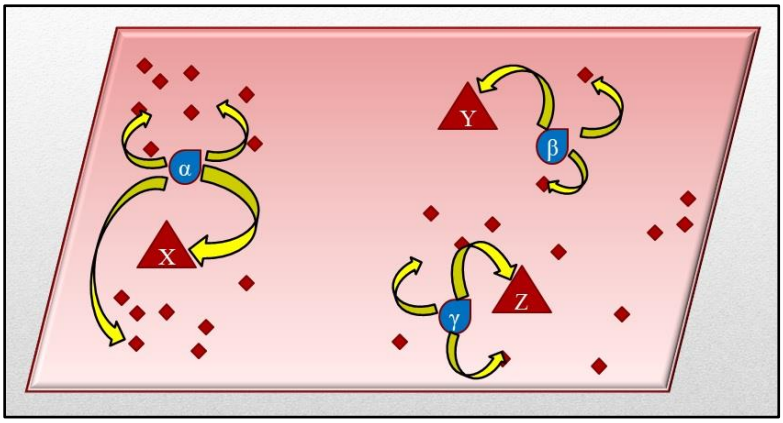

Fig. 2.1 Initial stage in a multi-agent system

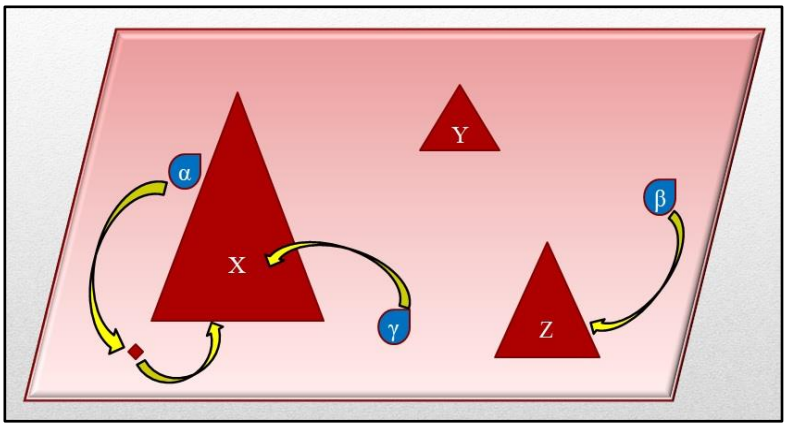

Fig. 2.2 Emergence of a new rule within the work-space

Following Kirsh (1996, p. 417), an agent trying to accomplish a task can redesign it acting *on the very topology* of the task-environment[7]. The physical modification of the space entails an easier solving process for the agent which provokes it but, being that physical space shared with other agents, this action affects the task-environments of those agents eventually *shaping their behaviour*. All these mechanisms can work because the environment is rich in structural information which heads agents' work in task-solving activity: distances, textures, volumes, colours and shapes are just some of the parameters which affect task-solving. This structural information provokes the emergence of a set of *implicit procedural rules* (namely, instructions to carry out a certain task) in a changing environment; then, these procedural rules feed back into the agents' behaviour.

To explain how these procedural rules may emerge I propose the following abstraction: let's think in a habitat as the work-space of a multi-agent system (Ferber 1995); a multi-agent system is a simplified artificial environment in which a group of relatively simple agents have some task to accomplish; they operate within a work-space, which is then an oversimplified artificial habitat, where they have to realize some

[7] A task-environment is an abstraction which designates the physical space where a determinate agent has to accomplish a certain task, including all the constraints, obstacles and resources that might affect the accomplishment of such a task.

tasks and they can manipulate their local environment to reach their aim. The programmer gives the agents a 'genetic program', namely, a set of basic rules they have to follow; for instance, a programmer could project a group of agents to gather some scattered blocks into heaps Fig. 2.1 as follows: 1) "Pick up a block when you see one"; "Drop it off when you find another block"; "If you do not find another block in *n*-minutes, drop it off"; "A block is the smallest element in the work space". After a while, a fifth new rule will emerge into the system: "Gather the blocks in the X heap" Fig. 2.2. *This rule is implicit in the changing structure of the work-space*; it is an outcome of the local actions of many individual agents within their habitat and, once it has emerged, it redesigns the environment itself eventually affecting the agents' future behaviour affording new environmental stimuli.

This clearly is a very synthetic description of these multi-agent dynamics, but it is sufficient to understand why Gordon (2015) uses the example of the ant colony to explain the emergence of collective intelligence (see *supra*, II, § 1.1): the X heap becomes an *attractor* because of its gradually acquired *relevance* within the work-space implied by the new topology; in the ant colony case, we find that foraging activity is guided by pheromones intensity in strategic places which modifies the topology of the habitat opening new foraging trails. The more relevant the information is, the higher is the number of agents which are attracted.

In 1959, the French zoologist Pierre-Paul Grassé published an article whose purpose was to explain how very limited individuals such as termites were able to coordinate and to build very elaborate structures for their nests (Grassé 1959). He concluded that each time that a certain termite α was starting a new pillar construction, the changes it applied at the habitat stimulated determinate reactions from the other termites of the colony, guiding their work. He called this principle of organization *stigmergy* (from Greek *stigma-* meaning 'stimulus' and *-érgon* meaning 'work'). Stigmergy is a very common mechanism in nature (see *infra*, V, § 2) and it is ultimately responsible for the indirect interactions among the diverse agents into a certain habitat (what the beaver is producing is actually an instance of stigmergy). Certainly, the pheromones dropped off by forager ants constitute a change of the topology of the habitat, being a

trace to follow which stimulates and informs the activity of the rest of the agents of the system. So we can say that stigmergy is the basic dynamics which allows the emergence of hetero-directed coordination into groups of individuals, *mediated by traces*. Though this is a kind of very general definition of stigmergy, I am using it here just provisionally.

Then, we have now a sketch of how some structures of information can emerge into an ecological space and how these structures can yield the emergence of a collective intelligence in a multi-agent system, such as an ant colony. Nevertheless, there is an important difference between stigmergy as the dynamics operating in a habitat in terms of the indirect relation among the diverse competitors acting in that particular space and, on the other hand, stigmergy as the coordination principle of the integrants of a collective intelligence: in the former case stigmergy results in an *adaptive organization* of an environment where different agents compete, like in the case of self-organization in the stock market (see *supra*, II, § 1.1.1); instead, in the latter case, stigmergy yields a genuine instance of *collective intelligence*, like in the case of an ant colony, sharing all the members of the colony the same genetic information, the same 'language' and the same purpose of survival in a habitat they share with many other competitors.

2.1.2. Public representations: stigmergy in collective minds

So far this framework works pretty well explaining the emergence of many instances of collective intelligence, both the (relatively) simple case of flock behaviour and the more complex case of self-organization in superorganisms like social insects. But the hypothesis I am proposing here has the purpose to explain the emergence of a *collective mind*. So, we have to reconsider here the case of an ant colony and the one of a honeybee hive: apparently, we are here in front of two specular instances of swarm intelligence (Trianni et al. 2011); however, some important differences still subsist and make the former one a mere case of collective intelligence, while some peculiar traits distinguish the latter one as a case of collective mind.

For what concerns the ant colony case, it is more evident in which sense stigmergy operates as a coordination principle: individual ants are phylogenetically 'programmed' (to borrow a term from artificial multi-agents systems field) to react to pheromone concentration in their habitat as a mark of belonging to the colony (ants from different colonies 'smell' different!); so that chemical signal is an attractor of ants *per se* and the instruction "pheromone → follow the track" is just one step of a *phylogenetically evolved algorithm*. In the same sense, when a forager finds a food source, it releases pheromone in that very point and all the way back to the nest because it is applying a further step of the algorithm: "food source → pheromone release". These two steps are actually structural changes which each forager introduces into the habitat, yielding indirect signals which shape its fellows' reactions and then provoke an avalanche effect. All these dynamics are, strictly speaking, *mechanical*: that is, we have no reason, in principle, to postulate any kind of representation/interpretation dynamics to explain the behaviour of a forager[8]; we could explain her behaviour as a simple reaction to a determinate stimulus[9].

This situation is slightly different with respect to the behaviour of the integrants of a honeybee hive: in fact, as I have already clarified (see *supra*, II, § 1.2.2), honeybees do employ a *representational code* for communication: each time a component of the hive discovers a new food source, it communicates to its fellows the exact position of the source in a range up to 10Km by means of the waggle-dance map (Seeley, Kirk Visscher, and Passino 2006). In this case, we have a (minimal) collective mind self-organizing for food collection. But still we have an act of communication which is possible through

[8] The long-standing discussion about the concept of 'representation' is wide and complex; however, analysing it exceeds the range of this thesis. I'm using the term 'representation', here, to mean any kind of *alleged* intentional content that might be conceived as an intermediate step between the perception of a concrete stimulus in the world and the production of an adequate response to it by a certain agent. There is, indeed, manipulation of a certain information in the case of an ant colony. However, individual agent's responses are the result of a mechanical application of a concrete algorithm: the single ant does not perceive the pheromone as a *symbol* for food; it just follows the pheromone to a certain destination.

[9] Still, this point is object of discussion. Some experts like Hölldobler and Wilson (2009) claim that ants can develop a *symbolic* communication: for instance, when the nest is under attack, ants release little quantities of formic acid (as they do with pheromones on foraging paths) to recruit soldiers to the place of the attack. Hölldobler and Wilson argue that the formic acid, when released in little quantities, is not useful at all for defence; instead, a symbolic association between 'formic acid' and 'enemy' developed as a phylogenetic trait from the *defence algorithm* "enemy → formic acid" into the *recruitment algorithm* "formic acid → enemy".

some changes introduced into a physical space by means of which an agent can indirectly *shape the behaviour* of his comrades: the map to reach the food source is 'drawn' into the hive through the iconic movements of the waggle dance as in a mime (Celli 2008, pp. 30-32); it is possible that many scout bees come back to the hive and communicate at the same time different food sources; when one of them reaches (for stochastic reasons) the 'attention' of a critical number of bees, the entire swarm opts for its map and follows it (Hölldobler and Wilson 2009, chap. 6; Seeley, Kirk Visscher, and Passino 2006, pp. 222-224).

So, my point here is that even publicly shared representations constitute a part of the structural information of the habitat which, in form of stimuli or constraints, redesigns the behaviour of concrete agents: in fact, the public performance of a forager bee is a spatial element deployed into the hive which offers a precise iconic information to her fellow bees about food location (Celli 2008, p. 35). In the next section I shall apply this framework to explain how publicly shared representations redesign the environment of the components of a certain human group, shaping their behaviour. By the way, as noticed by Huebner, publicly shared representations in a human collective mind are generally carried by concrete artefacts (e.g. public notes).

2.1.3. Public representations in the ecological niche: artefacts and stigmergy in the human case, from Otto's notebook to public lists

It is now necessary a jump back to the theory of the extended mind and its copious versions. In the first classic and prototypical case of extended mind, we have a subject called Otto who tries to obviate a memory deficit relying on a notebook which he continuously fills with important information. This notebook works for him as a *cognitive artefact* (Norman 1991) which subrogates his damaged biological memory. But, externalizing a cognitive activity into the surrounding environment through an artefact exposes it to multiple manipulations as it occurs in any physical medium shared by numerous agents. Otto's notebook is an instance of what has been called an *exogram* (Donald 1991, 2010), that is an external record of memory (see *infra*, IV, § 2). This is

important because Clark and Chalmers's seminal work approach entailed the externalization of mental processes into the surrounding environment *through the use of artefacts*. Indeed, although it was not then the main concern for Clark and Chalmers, we have to consider that the external environment the mind extends into is, after all, a physical space accessible to a wide range of other agents, an *ecological space* shared with many other actors. So employing an artefact in a cognitive task is a way to redesign and to make more comfortable for the agent a particular task-environment, which in Otto's case consists in remembering some pieces of information. But, as I have already remarked, a task-environment is an abstraction which defines the context where a determinate task takes place, including all resources and constraints which help or limit the agent's work. When Otto extends his memory into his environment by means of an artefact (namely, the notebook), he is 'interfering' with the task-environments of many other agents who share the same physical space as Otto, the same 'habitat'[10]. Then, as in the case of the beaver working to build its own ecological niche into its habitat, we find that the dam alters the topology adding new spatial information. So, any artefact shared into the habitat entails a change in the topology of that space. This is particularly true for what concerns exograms.

As I shall explain in more detail in the next chapters, there are many kinds of exograms: Otto's notebook, but any similar kind of written records, lists of items, sculpted glyphs and painted walls. In some sense, any kind of public representation might constitute an instance of exogram in so far as it is recorded in an artefact, and its durability depends on the durability of that very artefact.

At a first glance, the comparison between the beaver dam and a human artefact might rise some doubts, but what is relevant for my argument is that both of them are *alterations of the environment* which interfere with the normal behaviour of any other agent who bumps into them. These alterations are the traces left behind by any agent working in its habitat, which characterize the dynamics of stigmergy. These *traces* an agent leaves are eventually *signs* for the others who share the same environment. They

[10] I am developing, here, an implication of Clark and Chalmers's original example that they do not consider.

are traces for the agent who produces them because an agent leaves them *unintentionally*; they are signs for the other agents who bump into them because they carry information for them. And a sign is, by definition, an *information bearer*, because once it is created and shared, *it persists independently from its creator* as something available for the other agents into the system (Ricci et al. 2007, p. 133). This is evident with any kind of exogram: for instance, Otto might forget his notebook on the bus and someone could recollect it from the seat where Otto unintentionally left it, then she could have a look at the notebook and access at Otto's 'memory'; she could discover Otto's address and get there to give back the notebook to its troubled owner. A similar case would be the following one: Jane and Jack are a young couple living together in a cosy apartment; at least once a week they do a grocery shopping, but they are both a kind of forgetful people so they generally write down a grocery list before going to the shop. On a certain day Jane should go for the grocery shopping so she writes down a list of the things they need, but she forgot something important in her office that morning so she cannot go for shopping, then Jack takes the grocery list and goes to the shop; Jack does not know what they need, but he can rely on the grocery list to accomplish this task. This last example is not so different from any case of *transactive memory* (see *supra*, II, § 1.2.2).

This means that the information encapsulated in a sign is in some sense *quiescent* when it is not perceived by an interpreter, but it remains still *available* to be used by any agent *able to decode it*. I am still talking here of that structural information which is deposited into the system (the habitat with its inhabitants); this is slightly different with respect to the theory of affordances (Gibson 1979), as I shall explain later (see *infra*, V).

Both the beaver dam and an exogram are signs into a determinate habitat; the difference is that in the former case the habitat is a bare ecological space, while an exogram is shared into a cultural space. As I shall clarify later in more details (see *infra*, III), just as a beaver builds its own ecological niche into the habitat it lives within, we humans also build a *cultural niche* (Laland and O'Brien 2011) in addition to an ecological niche. This means that each human tries to build all around her *a culturally comfortable*

space, that is her own cultural niche. But we humans are social animals, this is why we generally collectively build a common ecological niche (as any other social animal species does) and a common cultural niche. Like in other social species, also in our case living in groups is a phylogenetic adaptation which implied many advantages in the perspective of survival. So each group features many agents who operate in the same physical and cultural space. As I remarked introducing the concept of stigmergy, when many agents do things in a shared environment, each time they do anything they leave a trace of their action, being this trace an alteration of the habitat; this alteration of the habitat changes the structural information available in the very habitat in terms of constraints and opportunities for action; in this sense that particular trace becomes a sign for the other agents who participate of the same habitat. But, if the ecological niche is made of things more similar to the beaver dam, a cultural niche is instead made of things more similar to Otto's notebook and Jane and Jack's grocery list; for instance, any kind of public representation, which indeed includes a painting or any other kind of public visual artefact. What I mean is that a cultural niche is in some sense the space of possibilities of what we usually call *collective imagery* and *public narrative*.

So I claim that the very structure of a determinate cultural niche where a human group fits, emerges from the local actions of single agents following a bottom-up dynamics. Nevertheless, once this structure has emerged, it establishes a certain range of constraints and opportunities (just as any ecological structure) for the future of the agents who inhabit it: the emergent procedural rules in this framework are *models* which exert a *normative power* as a top-down dynamics which feeds back to the individual agents of the system (see *infra*, V, § 3), providing (through constraints and opportunities) *instructions* about how to act, to think, to represent and, ultimately, to categorize some concrete contents. I shall then appeal to the folk-psychological theory of *mindshaping* (Mameli 2001) to explain how an emergent model can shape and therefore *direct* the cognitive behaviour of the integrants of a collective system, whom eventually this model coordinates into a collective intelligence. I shall argue this through the analysis of a specific case of study: the public use of images into a cultural niche and the emergence of representational canons (see *infra*, VI).

2.2 A brief sketch of my argument

Hence, I can summarize my argument for the culture of a concrete human group as its own collective mind as follows: the first criterion to distinguish a 'minded' entity from an merely 'intelligent' entity is that in the former case cognition consists in a meaning-making activity, while in the latter case cognition consists in adequate responses to environmental stimuli; the second step is that the meaning produced by a minded entity crystallize in public *representations*, such as the honeybee 'choreograms' (Celli 2008, pp. 30-35) or human exograms (Donald 1991, 2010); then, these representations imply the emergence of corresponding *normative practices* which regulate the agents' behaviour and the public use of such representations (e.g. a *code* implicit in canonical representations); consequently, normativity is a constitutive trait of the mental; eventually, in cultural niches, where manipulation of public representation occurs according to normative practices, there are the sufficient conditions for the emergence of a collective mind.

A final remark is needed about a criterion to discriminate two different collective minds: what distinguishes a determinate culture (namely, this specific kind of human collective mind) from a different one? I shall respond that it is the very *code* its integrants share. This is not simply the language (considered as a verbal communication tool) that they share, but the entire system of references they share when building their collective imagery, considering that *perception is multimodal and expression is multimedial*. So a code is the 'genetic' information identifying a certain culture just as the program is the 'genetic' information which distinguishes a determinate multi-agent system. The main particular trait of such a collective mind is that a human group can redesign its phylogenetically inherited information through the ontogenetic adaptation enabled by representations.

So far my proposal fits in Huebner's requirement of a collective production of representations and a collective use of them. In the next section I shall highlight what *novel properties* emerge with culture as a collective mind and what *functional gains* they entail for the members of the group.

3. Culture as a collective mind

The idea of considering the very culture of a human group as its collective mind is not completely new: there are some precedents of this concept both in contemporary philosophy and anthropology. Many anthropologists of the 20th century endorsed the idea that culture is a complex structure of behavioural schemes: Ward Goodenough argued that the culture of a society consists in all those things that someone has to know or to believe to be accepted by the members of that community (Geertz 1973, chap. 1, § III), defending an operational conception of culture as that practical information needed to act fairly within a certain cultural niche, while Roger Keesing literally defined the culture of a human group as a «superbrain that enables humans to solve survival problems in a wide range of environments» (Keesing 1974, p. 91). So in this framework culture appears to be a set of adaptively significant informational structures: this means that cultural adaptations fed back even to the biological structure of the human being, boosting his evolution (Geertz 1973, chap. 3, § II). Developing tools like knives improved manipulation abilities and erect position; cooking foods permitted a simplification in the digestive system which ultimately made possible for human organism to invert more energy in the brain (Aiello and Wheeler 1995; Aiello and Wells 2002); developing rituals improved the cohesion in groups boosting social activities. This results in that «a human brain to work (and even to exist) needs something external: it needs a reality – the culture – which, from the outside, conditions and permeates it reducing its possibilities» (Remotti 2011, p. 199)[11]. The famous example of using stars to orient navigation studied by Edwin Hutchins (1995, 2011, 2014) is another instance of how much pervasive culture is into our cognitive processes. But culture is not just a system of techniques which improve our practical life; in fact, culture includes communication systems and the very thought development which is eventually the fruit of a collective interaction (Donald 2017). This is because culture does not exist in individual minds, but out of them; in fact *culture is public because signification is public* (Geertz 1973, chap. 1, § III).

[11] Translation from Italian to English is mine, here and where not differently indicated.

Meanings are never casted by individuals, instead the production of meaning is a participatory construction, always moulded by supra-individual norms and institutional practices (Gallagher 2013, p. 8). Therefore, considering that in minded entities cognition consists in meaning production, and that this one is always social, cognition results to be something intrinsically social. So, being meaning making the ultimate essence of cognition, cognition itself has to be understood as *a culturally oriented collective endeavour*.

Moreover, the image of culture as an external structure of information which coordinates from the outside a group of agents has been the linchpin for a (then) new and brave proposal about human collective intelligence: the French philosopher Pierre Lévy proposed more than twenty years ago that the infosphere and the cyberspace enable us humans to cooperate in a highly interconnected way so that we can produce a collective intelligence by means of informatics and the Internet. He literally said that «the communicating informatics would manifest itself as the technical infrastructure of the collective brain or the *hypercortex* of the living communities» (Lévy 1997, p. 25)[12], while the Internet jointly with the totality of our technologies is considered by some theorists as the 'neural' ground of a real *global brain* which should integrate the whole humanity in a unique cognitive entity (Heylighen and Lenartowicz 2017).

Although these theorists presuppose the argument of shared information among the numerous integrants of such a massive cognitive entity, they seem more interested in the kind of *technology* supporting collective cognition than in the very kind of *cognition* that these technologies make possible.

3.1. The novel features of a collective mind

My proposal about the cultural structure as the collective mind of a human group develops instead in a different direction. Following the criteria of Huebner, Gallagher and Theiner and colleagues I exposed before, I now shall focus on those emergent novel features which are proper of the collective mind. So, while typical psychological features

[12] Translation from French to English is mine, here and where not differently indicated.

like memory or beliefs, in so far as they are meant in 'Cartesian' terms, appeal to the individual neural structure and dynamics, on the other hand collective memory and collective beliefs manifest very different and novel features which distinguish them from individual memory and beliefs. First, *collective memory*.

3.1.1. Collective memory

The well-known case of Otto's extended memory will work as a good starting point for my argument. Otto's memory extends into a simple artefact, a notebook; but what if Otto has a twin sister, Gina, whose memory is faltering as well? They could separately use their own notebook to take note of all their fleeting memories, but they are very united twins which spend almost all their time together; they share friends and plans, so they decide that it is not worth to use two different notebooks to keep trace of their chores; they will use just one! Would they share their memories? I would answer, yes. In fact, even if they used different symbols, labels and references to distinguish their peculiar memories in further details, this would be not so different from a case of *transactive memory* (maybe a bit muddled and slow instance of it, even though a functional one!). This is possible because their very memory does not consist in the artefact they employ; it rather consists in the very information encoded into that artefact. So, let's imagine a wider example of this kind of extended memory implementation: a work team has to accomplish every day a list of tasks; every night before, each of them writes down a task she thinks that has to be done on the following day on a keyboard they have in the operative room of the company, so every morning they all find a list with all the task of the day; each time an integrant of the team accomplishes a task, she simply checks it on the keyboard, so all the other workers will know which task has already been completed and which has not. So, being the information collectively shared, being the memory the very information, I would say that the memory itself is shared (through the use of the keyboard) by all the members of the group. The effects of keyboards and control panels on emergent coordination in groups are well-known (see *infra*, V, § 2.2; also Susi 2016), but these mechanisms have a

strong implication on collective memory: the keyboard works pretty well as a medium of coordination because it is a *vehicle* of information encoded in signs (words, images…) which are public, so *accessible* for all the integrants of the group. They are literally *representations* (e.g. of pending tasks) collectively built and collectively consumed by a group of co-workers.

But a keyboard is just a vehicle, then a *substrate* for annotations just as Otto's notebook. So, what if these very annotations were written down on diverse substrates such as walls or public monuments whose content was acknowledged and endorsed by the entire community and all its members could access to such annotations? Would these annotations constitute a common corpus of knowledge? So it seems (see also *infra*, VI). As the Russian psychologist Vygotsky claimed, memory is essentially an external activity, an activity mediated by signs: «[w]hen someone makes a knot in her handkerchief as a reminder, she is constructing the memorization process forcing an external object to remind her something, she transforms remembering in an external activity» (Vygotsky 1978/2018, p. 80).
And then:

> «The very essence of human memory consists in that human beings actively remember by means of signs. It could be said that the fundamental feature of human behaviour in general, is that human beings personally influence their relationship with the environment, and through that environment they personally change their behaviour, subjugating it to their control. It has been said that the essence of the very civilization consists in building monuments to never forget. Both in the knot and in the monument the most fundamental and peculiar aspect which distinguishes human memory from the animal one manifests itself» (*idem*).

It might be objected that this would not be the case of the majority of monuments because they were ordered by a determinate authority and this would violate Huebner's requirement of collective *production* (not just collective *use*) of representations but, even when an authority orders the diffusion of a message through its reproduction in public artefacts, the very content of that message (which ultimately is to be considered the

memory record) is something that appeals to a common ground of references – the collective imagery – which has gradually been built by the whole collective. When the Babylonian king Hammurabi ordered to sculpt and to display in public places his famous code, he also commissioned at the top of each stele a representation of the Babylonian God of Justice – Marduk – giving him the very code, so he was appealing to the collective imagery of his community to justify the prescriptive power of the code, because anyone in his community could recognize that character as a main piece of their pantheon – *their public narrative* – with all his peculiar features. I shall demonstrate through the case-study discussed in chapter VI that collective deposits of information emerge in reason of stochastic local alterations, and establish themselves as normative structures into a shared physical space as an outcome of the local activity of the members of a community; then, they shape the community itself. It is eventually this public narrative collectively built through the gradual contributions of the individual agents what constitutes the collective memory of a human group, and this public narrative has to be conceived as the totality of the information publicly shared by means of artefacts in a cultural niche, always accessible and modifiable by any of them.

Hence the most important point about collective memory is that memory records, being encapsulated into concrete artefacts, are ultimately exograms. This makes the emergent collective memory manifest some peculiar features which distinguish it from individual biological memory: the latter is on-line, directly connected to a determinate biological subject, neural-based, internal, private, synchronic and phenomenal; the former is instead off-line, disconnected of any particular biological subject, artefact-based, external, public, diachronic and symbolic. The second novel emergent property are *collective beliefs*.

3.1.2. Collective beliefs

Any act of reference expresses a belief[13]. When I say, pointing at a certain flower in my garden, "This is a rose!" I am actually meaning "I believe that this is a rose!". This seems almost trivial if said with respect to a bare natural object I experience in my environment. But I might also point at a photo saying the same utterance "This is a rose!". I am now expressing a belief with respect to an object which is not, literally speaking, a rose: it is rather an artefact which conveys *an iconic representation* of a rose. Now, any iconic representation is just a *translation into a graphic code* of the real object an agent wants to refer to (see *infra*, VI). Certainly a photo, as any other icon, reproduces just *some* of the features of its referent. It is true that a photo features a very high degree of iconicity with respect to the referent – it is for this reason that very few people familiarized with photography (and, clearly, with roses!) would disagree about this belief; nevertheless, the photo of a rose is not a bare copy of it. Just to make more explicit my point, a photo is a flat image and its three-dimensionality, perspective and depth are structures that we can 'read' in a certain photo only because we are familiarized with this medium.

These considerations about the graphic code would maybe sound more evident with respect to a drawing: I could draw a rose and then repeat the same utterance "This is a rose!". However, some people might disagree depending on my drawing ability or on their ability to 'read' a drawing. The point is that the act of reference mediated by a drawing has a very variable degree of iconicity; so, being iconicity a similarity relation with respect to *some selected properties* of the referent included in the graphic code, a drawing correctly refers to a certain object just for those people who consider as salient the same selected properties the drawing reproduces. This means that those people *share the same graphic code* that the author of the drawing. For instance a person living in the mid-19th might be familiarized with a well-known pre-Raphaelite painting representing

[13] Notice that I am not saying that any belief is an act of reference, but that any act of reference is a belief. Namely, *some* beliefs are acts of reference. Therefore, in this section my argument follows this scheme: every act of reference is a belief; some acts of reference are attributable to a collective subject; collective acts of reference are collective beliefs.

Ophelia, while she might not recognize women figures in the Picasso's *Demoiselles d'Avignon*. A person living during the Italian *Quattrocento* would surely not even recognize people in that painting. This is because they all live in diverse cultural niches where the graphic code acknowledged as the right one for representing that particular subject is different.

So, when I draw a rose and I say "This is a rose!", the graphic code I am employing expresses my belief of what a rose is (because the drawing refers to the rose). Consequently, my very belief (namely, my act of reference) *corresponds to the graphic code* I am using to depict the rose. Let's make it clearer: I am committing here to a teleosemantic theory of the intentional content (Millikan 1984), that is, the meaning of something is its proper function; a drawing is an iconic artefact which *refers to* a certain object (e.g. a rose); the meaning of that drawing is the *graphically codified* function I assign to it; so, this graphic code iconically expresses my belief of what a rose is.

As I shall argue in chapter VI, a determinate graphic code can become a *canon* when determinate conditions occur into the cultural niche of a certain community, namely, when it is recognised and employed as the correct code by all the integrants of that concrete community. This implies that the graphically codified function is collectively assigned.

A canon is then a *stigmergic emergence* (see *infra*, V, § 2.2) which is not reducible to the actions of any concrete agent in the niche, and it represents the normative translation into a graphic code of a concrete referent. So a *canonical drawing* of a rose is the *collective belief* (the collective act of reference) expressing the utterance "This is a rose!"[14]. In the same way, canonical representations of Pegasus or the Virgin Mary are the expressions of the respective collective beliefs.

So a collective belief is an act of reference collectively built and accepted as true by a certain group of agents; it is a normative emergence (a canon) which ultimately

[14] Someone might object that the drawing works as an utterance, so it is not the belief itself, but a vehicle for the belief. However, admitting this position would entail postulating a certain 'content' which is carried by the drawing, considering this content as something different from the drawing itself (e.g. some psychological element, some Fodorian symbol…). But this conclusion would ultimately force a return to the obscure notion of the 'mark of the cognitive' (see *supra*, I, § 1.1). Instead, I defend that the content (the meaning) of a drawing, coincide with the its assigned function.

shapes the minds of those very agents, regulates their private referents and gives coherence to their actions in time, contributing to the constitution of a collective identity. Finally, existing just because publicly represented into the physical environment of the niche occupied by a certain community, a collective belief acts both synchronically and diachronically on the integrants of that group; it exists in the physical external space of artefacts which constitute the niche. In chapter VI the emergence process of such a canon will be explained with further details through the analysis of the case-study.

To conclude this chapter, I would shortly summarize the idea of collective mind I endorse in this work, whose details will be analysed in the following chapters: a group of human agents produces a cultural niche as a self-organized system through stigmergic processes; this system can be considered as a collective mind of that very group a relevant part of which is constituted by the physical environment where it emerges; this physical environment is where both collective memory and collective beliefs are realized.

In the next chapter I shall focus on the analysis of this physical environment.

CHAPTER III

The Theory of Niche Construction

In the last chapter I have introduced a hypothesis about collective mind conceived as the whole cultural structure characterizing a human group and supporting collective imagery and public narratives. My proposal focuses on the emergent and novel mental features typical of a collective and supported by the physical environment a community lives within, namely its ecological, social and cultural niche. Such emergent collective mental properties are collective memory and collective beliefs, implemented by the public use of shared representations into the group's niche. I have also talked about the central role of artefacts as supports for public representations.

In this chapter I shall detail why considering the ecological space a community lives within is so important to understand the emergence of a collective mind. I shall endorse a scaffolded conception of mind, but my conclusions will converge on the fact that external cognitive scaffolds do not merely complement reasoning in individual minds; they rather enable the emergence of semantics (conceived as a property of collectives) which eventually allows the development of collective memories, publicly stored in external devices, and collective beliefs, connoted as normative referential relations.

In the first section I shall introduce the Theory of Niche Construction, that has recently become a revolutionary paradigm to explain the evolution of species: as the reader will see, niche construction suggests a switch in the epistemological focus on the evolutionary process, because it ultimately entails that agents living into an ecological space are able to co-direct (Laland and O'Brien 2011) their own evolution through the gradual accumulation of local modifications introduced in their habitat.

I shall also detail which niche-construction dynamics take place within an ecological space; this will be useful to make clear how structural properties emerging in a certain environment can underpin the development of a feedback loop between agents and the physical space they inhabit. Then, I shall stress that agents operating in a

certain habitat, when constructing their own niche, they also build *artefacts* which improve their possibility of resource manipulation within their niche.

In the second section I shall deal with the issue of the emergence of a cultural niche. Actually, this is a characteristic linchpin of human evolution which ultimately marked all the development of our cognitive relationship with the environment we inhabit. Cultural niche is, eventually, the theoretical connector which enables to link the scaffolded mind approach to the debate about the extended mind: namely, human mind works as it works just because it co-opts tools and artefacts in its processes and it relies on technological, social or cultural scaffolds of the most diverse kinds, such as alphabets, mathematical notations, institutions and rituals.

In the last section I shall connect the scaffolded mind approach with the third-wave debate about the extended mind (see *supra*, I, § 3) and, ultimately, with my own proposal about culture as a collective mind, considering the cultural niche as the space where human agents gather in groups and constitute communities thanks to the implementation of new cognitive behaviours, enabled by the development of new techniques, new technologies and new artefacts.

1. General features of the Theory of Niche Construction

The Theory of Niche Construction tries to offer a consistent framework of the evolutional influence of organisms' impact in their environment (Odling-Smee, Laland, and Feldman 2003; Laland, Matthews, and Feldman 2016). In effect, it opens a new perspective on the evolutionary process with respect to classical Darwinism, arguing in favour of an interdependence relationship between a certain environment and the agents who inhabit it. This theory is receiving increasing attention in studies about evolution (Laland 2004; J. H. Holland 2006; Okasha 2006; Dunbar and Shultz 2007; Kearney and Porter 2009; Townsend Peterson et al. 2011; Sterelny 2007, 2010, 2011, 2012; Brock, O'Brien, and Bentley 2016). My presentation of its main features will be based mostly on the foundational work by Odling-Smee, Laland and Feldman (2003), since it constitutes the main starting point for other extant approaches.

It is well-known that organisms play a main role in evolution carrying their genes, which are sorted out because of the selective pressures of the environment these organisms live within. However, organisms play a second crucial role in the evolution which corresponds to their phenotypes: in fact, organisms interact with their environments, consume resources, produce detritus etc. Doing all these things, they impact on their habitats and modify some of the selective pressures which characterize them. This is *niche construction* (Odling-Smee, Laland, and Feldman 2003, p. 1).

So conceived, niche construction manifests four fundamental consequences: 1) *ecosystem engineering*, consisting in all those changes an organism can introduce in its ecosystem to control the flow of energy and matter within it; 2) *modification of selective environment*, both for oneself and for other species acting in the same habitat; 3) a modified selective environment features new selective pressures, which represent the *ecological inheritance* for the new generation; 4) niche construction provides a second process for the *dynamic adaptive match* between an organism and the environment it lives within. These consequences have three main theoretical implications: a) for evolution, involving the very phenotypes as a second selection factor with respect to genes; b) for ecology, considering that niche construction entails the coevolution of organisms living in the same habitat; c) for humans, because niche construction sheds a new light on cultural elements and their relationship with our evolutive process (Odling-Smee, Laland, and Feldman 2003, p. 3).

Let's have a closer look at the first consequence of niche construction.

Any ecosystem is rich in niche-constructing species, for instance, leaf-cutter ants developed as a phylogenetic trait an interesting form of 'agriculture' which permits them to cultivate a fungus they use as a food source; this allows their colonies to reach huge dimensions and to have a deep impact in the surrounding habitat: they move and oxygenate the soil, leaving a softer stratum which is easier to penetrate for the weak roots of young plants; they also enrich the soil with many nutrients and in this way they create the conditions for an easier reproduction of those plants (Odling-Smee, Laland, and Feldman 2003, p. 5). But many other species influence the flow of energy and matter within an ecosystem, even very simple organisms such as cyanobacteria living in

the desert: they secrete polysaccharides which increase the compactness of the sand creating a crust which then makes it possible for superior plants to germinate; those microorganisms also reduce the soil water absorption, generating pools in occasional pits which ultimately permit the germination of more trees and the gradual formation of an oasis (Odling-Smee, Laland, and Feldman 2003, p. 7). A third example is the one of some snails eroding rocks when eating lichens: they transform the rocks in soil, affording new possibilities for other species (*ibidem*, p. 8). All these are manifest cases of ecosystem engineering.

From this first ecological consequence derives the main evolutionary effect of the niche construction, namely, the second ecological consequence: the modification of selective pressures. In fact, if organisms can control the flow of energy and matter within a determinate ecosystem, they can also affect the selective pressures existing in their own habitat and for the other species they get in contact with. Changes in selective pressures depend on the durability of the niche-construction actions undertaken by certain agents, so only *persistent structural changes* in the ecosystem can affect the very selective pressure and then they can imply an evolutive output. This criterion of persistence can be satisfied through two different dynamics: a) the ontogenetic change produced by a certain generation G_0 in a determinate environment at time T_0 can be repeated in the same way by generation G_1 at time T_1 and by generation G_2 at time T_2, because they share the same genetic information transmitted along the flow of time and generations, a genetic information which entails the transmission of the same niche-constructing phenotypes and cause the gradual modification of a determinate set of selective pressures which feed back to the adaptability of the new generations; b) the second case is when a concrete natural selection parameter is changed by a determinate generation and persists through the life of many other generations, for instance ancestral cyanobacteria which produced a primitive quantity of oxygen by means of photosynthesis in the early phases of the life on Earth modified the atmosphere and conditioned the subsequent biological evolution: they permitted the development of forms of life based on aerobic respiration (Odling-Smee, Laland, and Feldman 2003, pp. 9-12).

The third main consequence of niche construction consists in a second form of inheritance with respect to the well-known genetic legacy: when some agents produce persistent structural changes into their niche, these very structural changes represent a form of *ecological inheritance* for the future generations living in a certain habitat. So, ecological inheritance consists in the transformed physical space a generation G_0 bequeaths to a generation G_1, with all the structural changes which it can afford: a burrow, an increased (or a reduced!) amount of resources, a contaminated environment because of an excessive production of detritus etc. (Odling-Smee, Laland, and Feldman 2003, p. 13). Ecological inheritance does not depend on the presence of any environmental replicators, conversely to what happens in genetic inheritance; instead, it depends on the durability of the structural changes operated by some ancestral organisms in the considered habitat. This makes the ecological inheritance different from the genetic one in two main respects: a) it does not literally constitute 'information' in the same sequential sense the genes do, it rather consists in specific selection pressures which then select for specific genes; b) while genetic inheritance is transmitted just one-directionally from parents to offspring, ecological inheritance is transmitted also inter-species and bi-directionally from parents to offspring and vice versa (Odling-Smee, Laland, and Feldman 2003, p. 15). Moreover, it can be transmitted both during the lifetime of a determinate organism or after that (because of the persistence of structural changes), this is because it is *mediated* by the very environment all these different agents live within. Then, any organism which shares the same selective environment with another one is potentially related to it in reason of the changes it introduces or introduced in the past (*ibidem*, p. 16).

The forth consequence – *dynamic adaptability* – entails a new perspective on the adaptive match between organisms and the environment: in the received view about the evolution of species the physical space where an organism lives is conceived as 'decoupled', having *ab ovo* some natural features which select for such or such genes (this is natural selection); but a species can also adapt the environment instead of itself (Kirsh 1996) changing the habitat through ecosystem engineering. So the adaptive match is not just a question of survival of a certain species featuring determinate genes which

accidentally meet some external conditions; it is also due to the action of those phenotypes which transform the environment to improve their possibilities of survival in it and, doing this, they change the very selection conditions which are now more alike with respect to the needs of those phenotypes. In this sense the adaptive match is a dynamic process, because selection conditions and phenotypes *coevolve*.

The four consequences I have just exposed are ecological consequences of the niche-construction activity. These consequences, however, have some important theoretical implications that now I am going to present.

So, the first theoretical implication of the Niche-Construction Theory is a novel view about evolution, which is now conceived as a dynamic process entailing a *feedback* loop between a concrete environment and the phenotypes (namely, the agents) living there (Odling-Smee, Laland, and Feldman 2003, p. 20). Also, ecological inheritance increases the adaptability of the offspring in an environment modified by their parents to meet their phenotypes needs: for instance, in the case of oviparous species, eggs are deposited in a friendly nursery environment which elevate the possibility of survival for offspring upon hatching (*idem*). Finally, Niche-Construction Theory includes acquired characters as playing a determinant role in the evolution of species: phenotypes are not conceived anymore as mere 'vehicles' for genes, rather they are active elements which can change the environmental selective pressures and, then, they can feed back to their own genes' selection. For example, Galápagos woodpecker finches which learned to use cactus spines to implement their peck, created a woodpecker-like niche which gave more advantages to finches with a more fitting peck for grabbing and using cactus spines, because the new niche structure changed the selection pressures in favour of a more fitting peck for the new hunting-behaviour learned by the finches (*ibidem*, pp. 21-22). In this way information acquired through an ontogenetic process, even though it is erased when a concrete animal dies, it is nonetheless transmitted because of the modified selective pressures featured in the niche which an ancestral generation bequeaths to the subsequent ones in form of an ecological inheritance.

The second implication of Niche-Construction Theory concerns a new perspective on problems typically analysed in ecological terms: for instance, two populations living

in the same habitat can feed back to each other through their respective niche-construction activities, reciprocally changing some environmental selective pressures. Then, *coevolution of species* is possible through niche construction; this entails that coevolution is frequently *indirect*, mediated by some persistent structural changes operated by the diverse agents living in the same habitat, which also involve the abiotic (namely, non-living) components of an ecosystem (Odling-Smee, Laland, and Feldman 2003, p. 24).

The third implication of Niche-Construction Theory is about the relationship between material culture (including techniques) and the species evolution. Learning a new technique may entail a radical change in the structure of the niche, then in the corresponding selective pressures which have a decisive effect on genes; this is evident in the example of the finches using cactus spines to hunt their preys. The cultural factor is obviously exponential in humans and a well-known example of how cultural elements in human niche construction had then an effect on human genetics is found in cattle domestication: it implied an easily available quantity of milk as a food source, which then selected for the presence of the enzyme lactase also in adult human population (Odling-Smee, Laland, and Feldman 2003, p. 27). So «[h]umans are not just passive vehicles of genes, they actively modify sources of natural selection in the environment» (*ibidem*, p. 28); they are relevant for their evolution because of their phenotype, and a main feature of the human phenotype is *culture*.

1.1 The main dynamics of niche construction

It looks pretty clear that niche construction is not a particular feature of just a few main species playing a dominant role in their habitat; rather, niche construction seems to be more a fact of life in general: any species consumes resources and produces some kind of 'waste', having then some impact on the surrounding environment (Odling-Smee, Laland, and Feldman 2003, p. 36). This is because niche construction is, after all, a question of new physical structures and new matter and energy distribution patterns emerging in a certain habitat.

In this framework, a niche has to be considered as «the sum of all the natural selection pressures to which the [considered] population is exposed» (*ibidem*, p. 40). This definition clearly includes both favourable and adverse selection pressures. So, because selection pressures are relative to concrete organisms, a niche is a virtual space emerging all around a certain individual (or group of individuals). Consequently, we have niche construction when «an organism modifies the feature-factor relationship between itself and its environment by actively changing one or more of the factors in its environment, either by physically perturbing factors at its current location in space and time, or relocating to a different space-time address, thereby exposing itself to different factors» (Odling-Smee, Laland, and Feldman 2003, p. 41). When an organism finds a determinate feature-factor in its environment previously built by an anterior generation, this constitutes what we have called ecological inheritance, a legacy which could have been bequeathed both by genetic or non-genetic ancestors to the heir-generation (*ibidem*, p. 42). This point is crucial, because it remarks that ecological inheritance refers to something *located in the environment*, generated by a determinate agent but nevertheless independent of it. For this reason, ecological inheritance does not need a genetic connection between two generations; this is because it is a kind of *mediated transmission process* in which the very environment – by reason of its physical modifications – plays a main role, connecting both synchronically and diachronically different generations and different species.

Organisms can change the environmental factors (and the relative selection pressures) by means of two strategies. The first one is *perturbation*: it refers to all those physical changes an organism can introduce into its habitat in form of causal impact, such as exploiting resources, secreting particular chemical substances or constructing artefacts; the second one is *relocation*: it consists in all the actions an organism undertakes to move in space, choosing both the direction and the time of its migration (Odling-Smee, Laland, and Feldman 2003, p. 44). In perturbation an organism does not move nor adapts itself; it rather changes the physical structure of the space it is living within and, at the same time, the selection pressures implied by that very structure. In relocation an organism does not try to modify its ecosystem to make it match its needs;

instead the organism moves – migrates – to a new habitat where it is more fitting with the local selection pressures. Obviously, both these strategies are involved in most of the instances of niche construction: when building new structures in their habitat (e.g. their burrows), organisms also meticulously choose the location of such a structure; in the same way, when migrating to a new ecosystem, organisms introduce structural changes in it and their absence also impacts on their old habitat, being organisms a 'structural component' of the environment too. In fact, both organisms' presence and their absence in a certain ecosystem affect the very structure of that ecosystem, implying different patterns of the flow of energy and matter within it.

Moreover, niche construction can follow two developmental directions: a) niche construction is defined as *inceptive* when the change starts from the very agent, which tries to reach a more convenient matching point with the environment; b) niche construction is defined as *counteractive* when some patterns in the environment change and the agent has to react in order to recover the match with its habitat (Odling-Smee, Laland, and Feldman 2003, pp. 45-46). An example of inceptive niche construction is when an organism starts producing waste into its ecosystem and this changes the features of the niche, modifying its selective pressures; instead, an example of counteractive niche construction is when some feature of the environment accidentally changes and the organism reacts to preserve its fitness, like bird migration as a response to temperature change.

The two niche construction strategies I mentioned before and these two developmental directions of niche construction I just described intersect orthogonally (Odling-Smee, Laland, and Feldman 2003, p. 46): they can be considered as the four dimensions of niche construction, like four Cartesian semi-axes. In fact, we can have *inceptive perturbation* like in the mentioned case of a species living in a habitat which starts producing waste, changing in this way the physical structure of the environment; we have a case of *inceptive relocation* when a species consumes all the food available in its habitat and then migrates to find a new ecosystem rich in food sources, like in the case of locust swarms; we have *counteractive perturbation* when an organism reacts to modified environmental parameters remaining in the same place, for example plants secreting

chemical agents to defend themselves from parasites or beavers re-building the dam damaged by the stream of the river; we have *counteractive relocation* each time that a change in temperature, for instance, makes a species migrate from its ecosystem to a more fitting one.

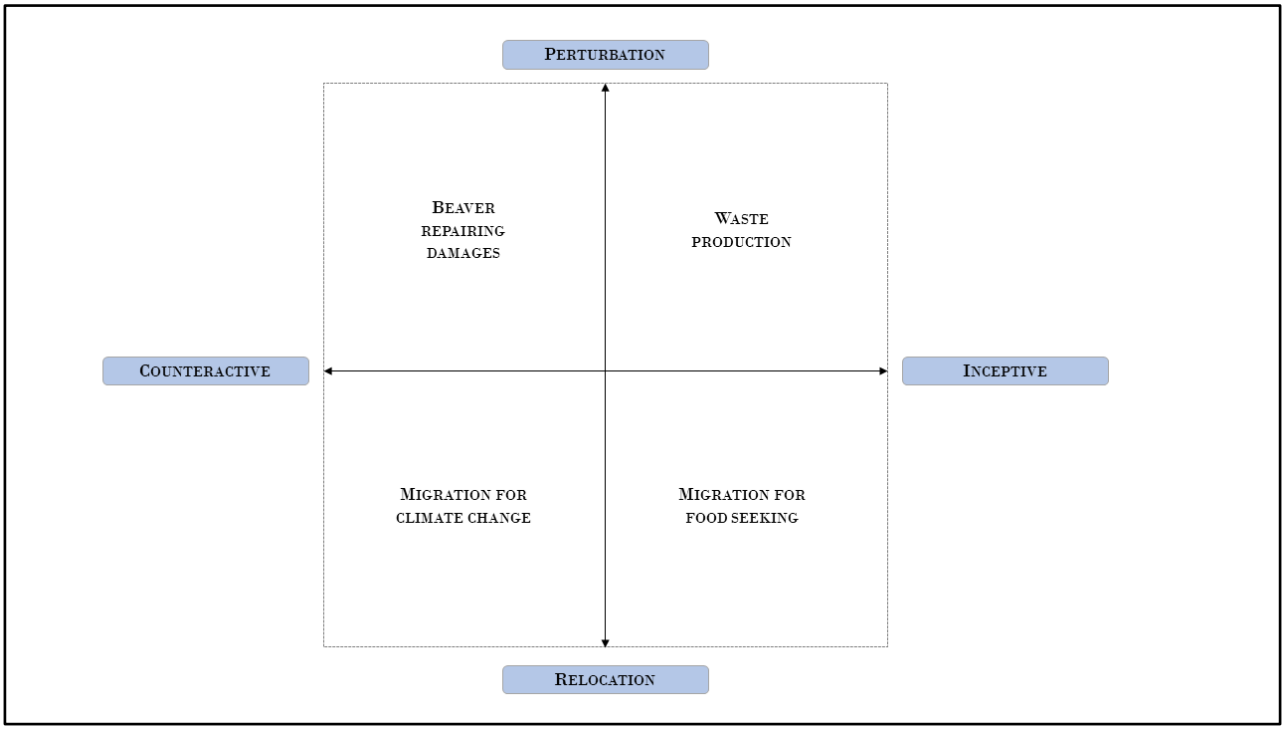

Fig. 3.1 The four dimensions of niche construction

Finally niche construction may have two different effects with respect to populations involved in it: a) niche construction can be *positive* when it increases the fitness of an organism into its ecosystem, for example when it results in an increment of food sources into the organism's habitat like in the case of leaf-cutter ants fungal agriculture, a case of ecosystem engineering which remarkably improves the survival chances of the colony; b) niche construction can be *negative* when its output reduces the fitness of the organisms involved, such as in the case of an exaggerated exploitation of food sources which reduce them dramatically, or in the case of an excessive production of detritus which ultimately results in the ecosystem pollution. So organisms co-direct their own evolution (Laland and O'Brien 2011; Kendal, Tehrani, and Odling-Smee 2011) through habitat manipulation both in ways that can increase or reduce their fitness, and this is up to the emergent physical structures in the environment which feed back to the

agents of the ecosystem, just as those emergent patterns which afford new procedural rules within a multi-agent system, re-organizing the behaviour of the agents and affecting the fitness of their responses (see *supra*, II, § 2.1).

1.2 Artefacts in the niche

Niche construction, especially when it occurs in terms of a perturbation dynamics, entails the production of *niche-constructed resources* (Odling-Smee, Laland, and Feldman 2003, p. 79). Niche-constructed resources are all the structures an organism builds into its habitat to increase its possibility of survival: for instance, burrows and nests are some of these structures. All these structures are *artefacts*, namely objects deliberately built by an agent to improve its life conditions into a certain habitat and they are maybe the best example of *inceptive perturbation*. When these structures evolve and are used beyond their original function to get a new objective, they can be considered as *tools*. For instance, many spiders evolved their original burrow structure transforming its dwelling in a real foraging tool: some species extend some trip lines out from their burrows; these lines work as an extension of the spider's sensory system, so when an insect touches one of them the spider gets out suddenly from its hollow and grabs its prey (Odling-Smee, Laland, and Feldman 2003, p. 85). So the burrow acquired during the spider's evolution process a new function as a hunting tool, a complex artefact the spider builds to enhance its foraging skills. The hollow it lives within is not just a dwelling now, but it constitutes also a main part of a complex trap: while the silk lines extend the spider's sensory system out of the burrow, the hollow hides the predator to its preys.

Then, the inceptive perturbation heading an organism to the construction of a burrow/nest subsequently feeds back to that very organism, pressing it to develop new behaviours, for example, in order to regulate temperature into its dwelling. The emergence of these new regulative behaviours affects the very structure of the nest: for instance, some termites model aeration holes in their pillars to manage nest cooling and to maintain a constant temperature convenient for the colony inside it (Odling-Smee, Laland, and Feldman 2003, pp. 87, 93); also honeybees developed a peculiar conduct for

thermic regulation of their hives: they drop some water on the nest to cool it by evaporation or rather they move their flight muscles to heat the hive (*ibidem*, p. 92).

Moreover, changes in the hive structure of some species of bees have induced an improved sociality within the swarm: removing the walls which separated cells in the hive created a common nursery area which ultimately promoted an increment of social life into the colony. The most remarkable point, here, is that the behavioural change (namely, an improved sociality) was eventually due to *a physical change* in the structure of the nest, which affected the relations among the members of the same colony: in fact, «in these bees the loss of the cell partitions is associated with a leap from solitary status to primitive eusociality» (Odling-Smee, Laland, and Feldman 2003, p. 101). All this is crucial to claim that the feedback loop which develops in any niche-construction dynamics is not merely due to the manipulation of some environmental parameters organisms can operate; instead, the very production and use of artefacts generates a feedback loop which ultimately stimulates some main behavioural changes in the very niche-constructing agents. All the resources – all the *artefacts* – organisms build when involved in niche construction are not trivial elements at all; rather these resources «constitute fundamental components of their worlds» (Odling-Smee, Laland, and Feldman 2003, p. 112). This is especially evident in human niche construction, where the use of artefacts is massive and the manipulation of the environment is filtered by cultural procedures. In the human niche, then, the very culture plays a main role in the feedback loop which the niche-construction dynamics generates: in human niche the selection pressures modified by the agents' action in their habitat are not mere environmental factors; instead, they are mostly emergent cultural structures of the niche which ultimately shape the behavioural patterns of the human agents living there.

2. The emergence of a cultural niche

As we have seen all along § 1, any organism tends to build its own niche in the ecosystem it inhabits, no matter how complex the organism is. In effect, niche construction is something which occurs independently of the complexity of the considered agent.

Nevertheless, when constructing their niche, human beings are special in some respects. In fact, they develop also *cultural structures* which result to be very powerful niche-constructing engines: «Cultural information, expressed in the use of tools, weapons, fire, cooking, symbols, language, agriculture, and trade, [...] played an important role in driving hominid evolution in general, and the evolution of the human brain in particular» (Odling-Smee, Laland, and Feldman 2003, p. 242; especially about the brain, see also Aiello and Wheeler 1995; Aiello and Wells 2002). So it sounds reasonable to say that human cultural activity is nothing different from a niche-construction strategy, just as cyanobacteria secreting polysaccharides, leaf-ants fungal agriculture or beavers building their dams (see *supra*, III, § 1): cultural activity is, after all, *a particular way to manipulate the environment* a human community lives within and, then, a way to change the selection pressures in that determinate habitat in a sense that may hopefully increase that very community's fitness. This is to remark that organisms co-direct their own evolution through niche construction (see *supra*, II, § 2.1.1 and III, § 1.1), and cultural behaviours in humans have the same role, namely the co-direction of human evolution. In fact, there is a close relationship between cultural processes and the very genetic evolution, namely, cultural processes can select for some peculiar genes instead of others, like in the case of cattle domestication (Cavalli-Sforza 2004).

Nonetheless, if in the received view about this close relation it has been argued that there is a *direct* correlation between culture and genes, niche-construction theorists claim that this correlation is instead mainly *indirect*: effectively, they suggest that it is mediated by niche-construction activity. This means that in the Niche Construction Theory, cultural practices can affect genes selection just because they alter the physical environment, so they modify the selection pressures and, through this modification, they ultimately influence genes selection.

Now, it is well-known that ontogenetic information cannot be directly transmitted through an immediate genetic mutation from parents to their offspring (Odling-Smee, Laland, and Feldman 2003, p. 243). The only way in which cultural processes may influence genetic evolution is by affecting the fitness of individual agents, increasing or decreasing their concrete survival possibilities, then also the possibilities a

certain genetic information has to be transmitted to an heir-generation. Considering this framework, cultural transmission systems provide human beings with an inheritance system carrying non-genetic information (*ibidem*, p. 248). In this way cultural inheritance permits the transmission of ontogenetically acquired information. In fact, if an inherited cultural structure persists through a sufficient number of generations, it might change the selective pressures for a determinate population and it might ultimately co-direct the genetic evolution of such a population. This is the so-called gene-culture co-evolutionary approach, which argues for a direct relationship between cultural and genetic evolution (Cavalli-Sforza and Feldman 1973a, 1973b; Feldman and Cavalli-Sforza 1976; Cavalli-Sforza and Feldman 1983; Feldman and Cavalli-Sforza 1984). I mentioned before the case of cattle domestication, and actually the inherited tradition of pastoralism is considered as one of the best examples of this relationship between culture and genetics: in fact, the establishment of this new cultural behaviour introduced durable changes into the human environment, which advantaged those individuals whose genetic mutations developed a greater tolerance for lactose in adult population (Odling-Smee, Laland, and Feldman 2003, p. 248). Here the direct relationship is possible because of the simple link between the usage of milk and its genetic consequences. However, the gene-culture co-evolutionary approach could become a kind of simplistic if generally applied to other instances of cultural manipulation of the habitat.

A case where the gene-culture co-evolutionary model does not offer a good explanation of the impact of a cultural practice on a genetic mutation is the one of the diffusion of a sickle-cell allele in some populations of West Africa, who ultimately manifest a reduced percentage of mortality due to malaria: in this area it is widely diffused the cultivation of some plants like yam and cassava; to grow these crops people living there cut clearings in the rainforests; this causes the increment of standing water, which eventually constitutes an excellent breeding ground for malaria-carrying mosquitos (Odling-Smee, Laland, and Feldman 2003, p. 251). Here the crucial variable is the amount of standing water, not the mere diffusion of a certain kind of crop, and it is an *ecological* element, not a *cultural* one. The increase of malaria mosquitos in this area

and the correspondent diffusion of malaria among the local population, ultimately selected genes in favour of a wider diffusion of sickle-cell allele (which makes cells more resistant to the action of the malaria parasite), but this selection is due to a change of the environmental pressures (in this case, a bigger amount of standing water) accidentally caused by a specific cultural behaviour (namely, the practice of a particular form of agriculture). So the relationship between the considered cultural practice and the final genetic effect is mediated by an *abiotic* element – the standing water – which is an ecological variable. This point supports the thesis that culture does not condition genetic evolution only through a direct relationship; instead, cultural practices *indirectly* affect our genetics in many situations because they are, after all, our peculiar human way to build our own niche and, as a consequence of our specifically human niche-construction activity, cultural practices can influence our genetic inheritance[1].

An important part of human niche-construction activity consists in the recollection and transmission of information. There are three levels of information-gaining-and-transmission process in humans: the genetic process, the ontogenetic process and the cultural process. The first one refers to all the useful mutations at a phenotypic level which are retained in the gene pool and transmitted to the offspring; this is a kind of phylogenetic learning and the most basic process responsible for niche construction (Odling-Smee, Laland, and Feldman 2003, p. 254). Ontogenetic process, instead, refers to all the instances of information gaining which are related to a particular individual; they can have diverse bases, for example the immune system or the nervous system in vertebrates: the immune system develops an idiosyncratic knowledge of pathogens, while behavioural routines are learned by means of a trial-error pattern by the nervous system to individuate new things or situations which might be positive or negative for the subject (*ibidem*, p. 256). This equips the subject for a quick and smooth adaptation to local conditions. However, this kind of ontogenetic information can be transmitted to the heir-generations only through niche construction, like in the case of woodpecker finches I mentioned before in this chapter: a new learned information gives an

[1] About the relationship between genes, culture and niche construction see also a more recent work of Michael O'Brien and Kevin Laland (2012).

evolutionary advantage to those finches that have a more fitting peck for grabbing and manipulating cactus spines; these finches presumably increment their possibilities of reproduction and, then, the frequency of their kind of peck in the offspring (*ibidem*, p. 258).

Cultural information transmission is in some respects special: it is multi-directionally transmittable among the members of a certain group; in fact, cultural schemes can pass from the elderlies to the youths or vice versa; they can also pass from a certain individual to a different one by imitation or social learning (Odling-Smee, Laland, and Feldman 2003, p. 260). Anyway, cultural information is never randomly transmitted; rather, smart behavioural patterns are selected in the cultural niche just as genetic traits are selected in an ecological niche (*ibidem*, p. 259). Effectively, cultural variants are generated by individuals as the outcome of a social learning and they are eventually selected because of a *differential adoption* in the considered population. In this sense, cultural processes can also accumulate functional responses with respect to the environment, just as it occurs in biological evolution (*ibidem*, p. 258). In this context, cultural learning reveals itself as a real shortcut with respect to the phylogenetic learning: it carries specific information about things which are useful or dangerous, about environmental resources of threats for the organism. What is peculiar of cultural information is that it requires non-genetic channels of communication and that it can be decomposed in 'chunks' like bits of information; this means that any individual can store some cultural knowledge and then she can transmit it in discrete units (Odling-Smee, Laland, and Feldman 2003, p. 259), like in the case of *artefacts* in general and *signs* in particular (see *infra*, V, § 3). In sum:

> «[c]ultural niche construction modifies selection not only at the genetic level, but also at the ontogenetic and cultural levels as well. By modifying the environment, niche construction creates artifacts and other ecologically inherited resources that not only act as sources of biological selection, but also facilitate learning and perhaps mediate cultural traditions» (*ibidem*, p. 261).

Then, human niche construction is mainly guided by cultural inheritance. Even though, the acquisition and transmission of such a cultural knowledge largely depend on the genetic inheritance system and on ontogenetic learning processes.

Niche-construction theorists claim that what distinguishes human cultural niche from many instances of animal proto-cultures (e.g. Galápagos woodpecker finches) is the fact that human cultural niche shows a goal-directed progress which does not exist in the animal case (Odling-Smee, Laland, and Feldman 2003, p. 261). What I shall argue all along the next chapters is that this is due to the human capacity for abstraction, which ultimately is the production of meaning. I shall argue that it consists in a skill related to the reference process and the emergence of semantics as a property of collectives, as I argued in the previous chapter (see *supra*, II, § 3). Ultimately, this goal-directed character in human niche construction is possible because of the emergence of a *symbolic culture* in the *Homo sapiens* niche (Sterelny 2011), which is a specific outcome of the social dynamics which gradually developed in the human niche during the evolution from hominins[2] to humans (Sterelny 2007, 2012).

As a matter of fact, social life became quite early a determinant feature of the first hominin communities: the new food sources and the new habitats these peculiar primates started to use and live in, made collective actions and social life a necessity[3]. This was likely because the new kind of alimentation these groups of hominins adopted entailed the implementation of new complex techniques, such as cooking, which on one hand made available new and more nutritive food sources but, on the other hand, these new alimentary habits also implied a division of labour which required an increment of social life (Sterelny 2007, p. 719). Nonetheless, cooperation is only durable when a fair division of benefits among the cooperating agents is realized (*ibidem*, p. 725).

---

[2] 'Hominin' is a technical term which generally refers to all the primates considered as ancestral to humans or related to them, indeed including humans themselves as well. Nevertheless, many biologists and paleoanthropologists (including Sterelny himself) use this term to refer only to the genera *Homo* and *Australopithecus* and their extinct relatives, thus using the word 'hominin' as a synonym of 'hominid'. This last is eventually the sense in which I, myself, am using this word here.

[3] About how the development of pro-social behaviour may positively influence the fitness of a group and become also a selective character, see a recent work of Paul Ryan, Simon Powers and Richard Watson (2016).

Then, social life can emerge when the task-environments (see *supra*, II, § 2.1) of many conspecifics are 'superimposable', so that the individual agents are not real competitors for resources; instead, the adoption of social practices results in a more convenient management of the work/benefits relationship and eventually in a common adaptive advantage (Nowak 2006). This is a well-known evolutive feature proper of many social species, especially social insects such as primitive solitary wasps, bees and ants that, building their niche, gradually developed eusociality (Hölldobler and Wilson 1994, 2009).

As I remarked in the previous section, ecological inheritance does not just refer to a set of artefacts or resources such as mere abiotic structures an ancestor-generation bequeaths to an heir-generation; rather, ecological inheritance also consists in *the totality of the agents* sharing the same niche a previous generation bequeaths to its successors (because agents are indeed material elements of the niche as well!). Let's think of an hypothetical example: in a certain species it occurs that some individuals have in their genetic pool some mutations which might make them have, as a *phylogenetic* trait, more pro-social attitudes (e.g. gathering together in the same places, eating together or sleeping together); the new offspring, then, are born in an environment which now features determinate social dynamics (e.g. conspecifics looking at the same food source or the same partner); some of these offspring develop *ontogenetic* adjustments trying to improve their fitness (e.g. strategies to reduce frictions with conspecifics, like social grooming); eventually, the most pro-social individuals increase their reproduction possibilities (Brent et al. 2013) and, then, they cast a pro-social ecological niche where pro-social attitudes are found in the environment where the heir-generation is born and these attitudes certainly are encouraged by the parents too. This is how a *social niche* might develop into an ecological niche (see also Powers et al. 2010; Laskowski and Pruitt 2014; Ryan, Powers, and Watson 2016)[4].

[4] Notice, by the way, that the social niche is nothing different from a set of social features of the ecological space the considered agents are sharing; clearly, it does not constitute a different ontological level at all. A similar consideration has to be done about the concept of cultural niche, which I am going to detail in the next pages: 'cultural niche' certainly refers to some features of a concrete ecological niche as well, which constitute an ecological inheritance whose effects have a cultural nature (e.g. *artefacts* and

This is clearly the case of hominins' social niche, where pro-social attitudes developed as an ecological inheritance. This 'ecological factor' presumably changed the selective pressures of these hominins in the sense that, for instance, new bigger animals could be hunted by the group and many other dangerous animals were not a threat anymore for that collective. This means that the ecological niche inherited by a particular generation of hominins was also a social niche, namely a niche characterized by mutual relations among the members of a certain group.

When, during the evolution, hominid groups became too much socially complex both vertically (hierarchical levels) and horizontally (individual diversity among the members of a same hierarchical level), the development of a *cultural code* became compulsory as well. This is because, the social niche has to be conceived, at last, as a set of social variables (namely, all the possible relations among the members of the same group); when these variables multiply, the emergence of some 'tools' to reduce ambiguity is to be expected: for instance, the establishment of defined social roles and 'institutions' (e.g. marriage, rites of passage at the adult life…) might have implied the correspondent development of 'reasons' to justify them and related 'norms' to support such hierarchy (e.g. narratives).

Also, changes in social factors within a certain community can ultimately redesign the cultural niche of that very community: for instance, social roles and the relative norms might be established with respect to those individuals who, for some reason, control a crucial amount of resources in the shared niche; then, those very individuals might also become *cultural* points of reference within their communities[5]. Symmetrically, changes in the cultural niche could affect the social niche as well, like in a feedback loop where influence is reciprocal among the terms of the relationship. Let's think, for instance, of the influence of social variables on marriage customs, which constitute a

---

technologies; *techniques* in general and *cultural practices* such as good manners or other peculiar customs which regulate social life for cultural reasons for a concrete group of agents).

[5] An illustrative example of this transition could be the emergence of a certain hierarchy in the archaic Greece: the *wanax* ('kings') were originally rural leaders whose power was based in the control of a certain amount of herds (Kilian 1988): the more the sheep, the more the social power of their owner.
What is particularly interesting, by the way, is that *wanax* were not mere 'political' figures: they were also the main ministers of religion and the 'highest judges', so also a cultural point of reference for the community.

cultural trait: family's wealth is a factor which, in some societies, likely provokes wife's moving to her husband's house or vice versa (Lipatov, Brown, and Feldman 2011). By contrast cultural changes can affect the social niche as well. Contemporary cases could be, for instance, the introduction of the *divorce law* or the *strike right* in societies where they did not exist before: in the first case the cultural change likely induced more flexible family structures, while in the second case the cultural change radically modified the previous work relations.

So, in an evolutionary focus, a more efficient communication in *Homo* groups might have resulted from the establishment of less ambiguous signals so that coordination could be managed through the development of norms and customs (Sterelny 2007, p. 721). The emergence of such norms might have eventually resulted in a transparent social environment.

This transparency certainly was an adaptive necessity: when a group of individuals starts cooperating, this requires from each of them a considerable energy inversion in social life which risks to be frustrated if it is not compensated by an adequate benefit. As some theorists have noted, collective action always hides the threat of a 'Machiavellian loop' (Sterelny 2007, p. 722), namely a certain member of the group could try to take advantage of the work of the others, enjoying benefits for a work she has not done. This framework is easily understandable within our contemporary anonymous mass-society, where we are reciprocally all strangers: we live in a society where personal relations are not durable and they often result to be single events in our life. But the evolutionary conditions in which human social interaction emerged were radically different; in fact, while in the contemporary mass-society individuals are highly mobile (because also their resources – especially money – are highly mobile), hominid ultra-sociality emerged instead in very small and stable groups (Sterelny 2007, p. 721), where inter-group shifting was a very unlikely possibility.

The point here is that the development of foraging habits entailing cooperative actions firstly posited the problem of *commitment* and *defection* in the group. In this framework the emergence of norms and customs is clearly justified by the adaptive necessity to disambiguate defection in cooperative social groups: «[n]orms make the

social environment more transparent» just in this sense (Sterelny 2007, p. 722). In a small environment with a long history of interaction the risk of defections is reduced, thus cooperation is stimulated by the very *demographic structure* of the group (if we are few, we all know each other). This is because in relatively small symbolically-marked populations the costs of cultural mimesis are too high to compensate possible defections: when a subject makes part of a very small community living in a huge hostile environment, if she decides to betray the confidence of her fellows, she will certainly be expulsed from that group; then, even if she found a new clan which accepts her, she should nevertheless invest a lot of energy in learning the specific customs of the new group (e.g. the peculiar 'communication code' of that community[6]).

So, the emergence of a *social* niche in the human ecological niche made ultimately inevitable the emergence of a human *cultural* niche as well (Laland 2017). This one implied cultural costs for defection because the group membership, when vertical and horizontal complexity are significant, results to be symbolically marked. Then, to get the group membership one has to learn this cultural code, which requires a considerable mental energy. On the other hand cooperation compensates the agents with many benefits both material (for instance new kinds of food source are now available) and informational (because acting in a group improves human skills for interpreting the intentions of conspecifics), encouraging pro-social behaviours (Sterelny 2007, p. 726). This affords the integrants of a group the access to much more information about the environment they live within, because human adaptation to cultural learning makes possible to use the minds of the others as information sources (*idem*). Effectively, when cooperating in a group, agents do not just learn things about their co-workers, they also get from these co-workers a lot of information about the opportunities afforded by their ecological niche. Then, hunting together provides the group members with both *ecological information* about resources and dangers available in the niche, and *behavioural information* about their fellows, testing their reactions in a cooperative environment (e.g. whether they are trustworthy or not).

[6] Just think of the common experience of the difficulty to learn a new language when you suddenly move to a new country!

So the niche-construction paradigm radically redefines the terms of the relationship between organisms action into their niche and the niche itself, conceived as the vehicle of natural selection. As I have so far remarked in this chapter, organisms are not simply moulded by the natural selection; instead, they actively modify the selective pressures in their environment, through their daily actions. Nonetheless, it is important to note that these organisms do not consciously design their environment pursuing the 'goal' of an improved fitness in it; they rather act locally and adaptation is just a possible consequence of their local actions: in fact, it is also possible that individual agents operating into the niche do not ultimately adapt the environment in any favourable way with respect to their genotype. The ecosystem engineering they undertake may actually have a notable impact on the structure, the composition and their variety of the community of organisms which inhabit the niche itself: in effect, young beavers do not just inherit a dam from their parents, but all the microorganisms, plants and animals living in there (Laland and O'Brien 2011, p. 192). Then, niche construction is «a two-way process» (*ibidem*, p. 193): organisms dealing with environmental problems to solve them, necessarily posit new problems for the next generations[7], for the other species sharing the niche and, certainly, for themselves as well. This is the sense in which it is to be understood that niche construction heads the agents' behaviour and that it features a continuously modified developmental environment. These standard dynamics of niche construction also characterize social and cultural niche for species featuring social (or cultural, in human case) behaviours, considering that in the social niche the landscape is not merely physical but it mainly consists of social relations and *behavioural elements*, while the landscape of a cultural niche is composed of *symbolic elements* which are mainly artefacts, norms and customs (Sterelny 2012, pp. 151-153). Obviously, both the social and the cultural niche occur into an ecological space; actually, the social and cultural niche could be conceived as some 'advanced dimensions' of the human ecological niche.

[7] Niche construction may be so *positive* as it may be *negative*, depending on the particular selective pressures it changes (see *supra*, III, § 1)

As I have just noted analysing the focus proposed by Kim Sterelny (2007, 2011, 2012), even though niche construction is a widely diffuse activity in nature, because even agents with limited intelligence build their niche, it is anyway evident that humans, along their evolution, adapted their ecological niche in a very special way: as many other animals, they developed pro-social habits which resulted in a structured social niche; but, as a consequence of their ultra-sociality (Sterelny 2007, p. 721), human groups selected for the production of systems of norms and customs which ultimately constitute the core of the cultural niche (Sterelny 2012, pp. 166-167).

This is a very important point because, as I highlighted at the beginning of this section, cultural processes are very strong tools for the construction of the human niche (Laland and O'Brien 2011, p. 194). Ontogenetic information acquired by the agents modifies the ecological niche in a manner that boosts the selection of particular learning strategies and eventually heads to the establishment of *behavioural traditions* (Laland and O'Brien 2011, p. 196). In fact, humans inherit semantic information (Odling-Smee, Laland, and Feldman 2003, p. 253) through the behaviour of both peers and elders conspecifics; as a consequence, human socially-built environments influence the behavioural development: for instance, *playing* and *teaching* offer learning scaffolds (Kendal, Tehrani, and Odling-Smee 2011, p. 787)[8]. Human niche construction implies, then, a remarkable use of artefacts which are at the same time a source of biological selection and a vehicle of cultural traditions. Eventually artefacts permit a durable abiotic transmission of such cultural traditions between different generations, allowing human gene-culture coevolution. Thus it appears that language, cooperation and cultural intelligence in general have gradually coevolved, through a process of accumulation, with those technologies and social conventions afforded by these cultural skills themselves (*ibidem*, p. 788).

[8] Vygotsky extensively talks in his works about the relevant cognitive role of playing and teaching as social activities which structure children's developmental environment, providing them with a guided learning path characterized by rules, examples and many other elements which contribute to the constitution of that 'proximal space' so necessary to make the developmental process affordable for the child (Vygotsky 1978/2018; Meini 2012; Krueger 2013).

2.1 Factors involved in cultural niche construction

An important step of cultural niche construction in human evolution is the one characterised by the advent of 'behavioural modernity'[9] in *Homo sapiens*, as Sterelny defines it, around fifty-thousand years ago (Sterelny 2011, p. 810).

Sterelny tries to explain the evolutive emergence of those practices responsible for cultural niche construction in humans. His central thesis about this evolutive process is based on *social learning* as the main feature of human cognition. Thus, social learning firstly depends both on the accuracy of cognitive mechanisms for learning and on the control agents have on their developmental environment; secondly, social learning depends on the deep changes in material culture which gradually emerged during the evolutive process (Sterelny 2011, p. 810). Considering this framework, human cultural life would ultimately depend, in Sterelny's argument, on our capacity to accumulate and to transmit *cognitive capital*. This transmission is possible by means of two main elements: on one hand, language and imitation considered as *transmission technologies*; on the other hand, *developmental environments* redesigned and adapted through niche construction.

Thus, developmental environments permit the constitution and accumulation of a certain cognitive capital, while transmission technologies allow an ancestor-generation to bequeath this capital to an heir-generation. The gradual constitution of a concrete cognitive capital into a specific developmental environment is possible through *social learning*, which is mainly realized in the 'apprentice mode': this consists in «learning by doing» (Sterelny 2011, p. 810) like in a *training*, a training which takes place into a determinate physical environment rich in informational resources[10]. This means that the apprentice learns by observing and imitating her 'teacher' acting in a *structured space* (Prinz 2002; Gattis, Bekkering, and Wohlschläger 2002; Bekkering 2002; Vogt 2002; Tversky, Bauer Morrison, and Zacks 2002; Reed 2002; Rizzolatti et al. 2002).

---

[9] Sterelny uses this label to refer to the advent of symbolic cultures and how this advent affected human behaviour, namely, with the emergence of the abstract thought.

[10] The reader should keep in mind, here, the Vygotskian concept of *proximal space* (Vygotsky 1978/2018, pp. 124-134; particularly, p. 133).

A certain space may be considered 'structured' by reason of a set of peculiar *artefacts* which shape the perception and the possible uses of the space for the agents involved in learning tasks; such a space can also be considered 'structured' because of the social influence of peers and elders following the learning process of an apprentice and heading it through *advices* which form a structured problem-solving path for the beginner. This structured space is, after all, always an *ecological space*; it is an ecological space that features cultural elements such as artefacts, technologies, techniques, norms and other practices. This is a cultural niche, namely, what ultimately allows the accumulation and transmission of cognitive capital through generations: in fact, cognitive capital is just a particular form of *ecological inheritance* (see *supra*, III, § 1).

We have archaeological evidences supporting that the advent of 'behavioural modernity' was likely around fifty-thousand years ago, at least one-hundred-thousand years later than the appearance of the biologically modern man. We can observe many traces indicating the existence of *symbolic cultures* in that phase of human evolution (Sterelny 2011, p. 810). Behavioural modernity is so important because it marked a radical change in the human social niche: with the advent of symbolic cultures, the members of a group stop just *belonging* to that group and start *identifying* themselves as members of such community. In this period agents start using determinate insignias as membership marks, this meaning that they were aware of their group and they identified themselves with their community. These sapiens became aware of their *collective identity* (*ibidem*, p. 812).

Cultural niche construction, in this context, should have gradually occurred in that relatively large lapse of time included between the advent of biological and behavioural modernity of our species. Organisms living in an ecological space coevolve with such a space (see *supra*, III, § 1), they enrich it with artefacts and change the range of stimuli the niche affords to its inhabitants. A social niche develops when conspecifics make part of the niche where an agent lives, so that even *social relations* contribute to the constitution of the ecological space where that concrete population lives. Considering this point, Sterelny connects the established use of some artefacts such as *material symbols* which improve coordination into the social niche as factors which changed the

selection pressures of that very niche, fostering the development of secondary uses of those very artefacts in a symbolic framework.

So, for instance, we have no reason to induce from single burials that a human group elaborated an entire symbolic system to explain the passage from life to death; initially it might have just meant that such human group used to consider their ancestors' corpses as something different from a mere refuse; still nothing implies that also something like a 'mythology' or a 'religion' was connected to such use (Sterelny 2011, p. 812). Nevertheless, the established custom to 'bury the dead' could disseminate burials in the ecological niche occupied by a certain population in such a way that specific places for burial could emerge and consequently acquire some symbolic value. Redesigning the ecological niche through social practices entailed the development of symbolic uses.

A second example to illustrate the relationship between ecological, social and cultural niche is offered by the diffusion of ochre body-painting. Ochre marks maybe started to diffuse as a hunting-camouflage technique which, once established, worked as an ecological inheritance supporting the emergence of a symbolic function, namely the membership of a concrete group (Sterelny 2011, p. 812). While the diffusion of this symbolic function of ochre body-painting might have been justified by the necessity of long-distance inter-groups communication, the relatively more recent use of primitive jewels made up of shells and beads might have worked instead as an ecological inheritance which supported the development of an intra-group communication system identifying rank, role and status of a determinate individual within the hierarchy of her group.

Sterelny explains the emergence of this derived symbolic feature with the *demographic increase* in groups: in so far as material symbols (including ochre painting and shells jewels) work as *signs* (then, as information bearers) they are superfluous in a simple social environment; instead, when social environment gets denser, insignias become crucial for recognition, for instance to identify (and to be identifiable for) an individual who is not directly known, localizing her in the social network of the community (Sterelny 2011, p. 813).

Building a cultural niche entails that humans do not move just in a 'material' environment, but they also build and navigate an 'informational' environment: that is, they can now manipulate elements which do not carry merely ecological information, but a representational one through the use of signs[11]. Therefore, the use of material symbols implemented the possibility to create *symbolic schemas*, such as systems of numerical notation which enabled humans to think of quantities in very novel ways (Sterelny 2011, p. 813); material symbols also afforded *a reliable support* for intergenerational transmission of information.

All this represents a remarkable ecological inheritance which permitted a collective management of information, because the cognitive capital was then reliably transmitted both vertically and horizontally among individuals of the same generation or belonging to different generations. So *recollection* and *reliable transmission* of cognitive capital are two preliminary conditions for the advent of behavioural modernity. When a demographic threshold is reached, a cultural niche emerges and collective management of information is possible (Sterelny 2011, p. 814).

Finally, the demographic factor also affects skill specialisation within the community, which eventually boosts the capacity for information accumulation. In this framework *redundancy* plays a central role: in fact, it completes the lacunae of low fidelity in information transmission, reproducing and spreading many 'copies' of a certain information available in the niche a determinate group inhabits. This is why the demographic factor results so important in the cultural breakthrough characterizing the behavioural modernity (Sterelny 2011, p. 817).

3. Cultural niche and collective mind

Niche construction theory, as I have been highlighting so far, introduces a very novel perspective in the analysis of the relationship between an agent and the physical space

[11] The information available into the human niche is not merely ecological. As the reader will see in more detail in chapter V, § 1, the physical structure of the human niche features ecological information in terms of *affordances*, while, when we humans manipulate signs as material representations, the available information is representational: namely, it consists in *constraints* and *conventions*.

she inhabits. In fact, while in the received view about organisms' evolution these ones were treated as individuals passively exposed to the forces of nature and to its selective pressures conceived as an independent environmental factor, niche construction approach rather remarks the mutual redesign operated by the environment and the creatures living in it.

Moreover, as put in evidence by Sterelny, the ecological niche an agent lives within does not just condition her daily life in individual subsistence operations; instead, in a certain ecological niche, some *structural social elements* may emerge as central features of the landscape an agent experiences where she is born and all along her life. These constitute what has been called 'social niche'.

Such structural social properties are an actual *scaffold* for the agent's life just as the physical surrounding space: like distances, shapes, textures, dangers or resources, conspecifics' behaviours all around us are crucial stimuli which invite us to act in such or such way. When a social niche emerges in the evolution of a certain species so that a determinate number of conspecifics start implementing cooperative dynamics (Nowak 2006; Powers et al. 2010; Ryan, Powers, and Watson 2016), these very social dynamics enable the group to develop *cultural features* which ultimately constitute the group's cultural niche (e.g. symbolically marked membership of the group).

The cultural niche approach affords, then, a useful heuristic tool to explain social learning: the apprenticeship of a beginner group member, in fact, takes place in a culturally engineered space which features particular conceptual scaffolds and is structured by the skilled teacher in such a way that makes it heading the learning process of the apprentice. This makes Sterelny affirm that cognition is a clearly social process in humans.

Considering this point of view, Sterelny criticized a few years ago (2010) the extended mind thesis, contesting the explanatory efficacy of Clark and Chalmers's theoretical framework. In fact, he argued that their argument would work at most as a special case of human niche construction, in terms of scaffolding the cognitive environment of an agent by means of determinate artefacts supporting some specific cognitive dynamics; in their particular case, some mental properties like memory and

beliefs (Sterelny 2010, p. 466). Sterelny's point is that, as many other species, humans also rely on their surrounding environment to accomplish a lot of tasks in their daily life (Kirsh 1995). He remarks an important point about how our digestion process changed along our biological and cultural evolution: the implementation of cooking techniques made possible for humans to access many new sources of food which were not available before; also the implementation of such technology made possible for these humans to develop shorter and more simple guts with respect to their more primitive ancestors (Aiello and Wheeler 1995; Aiello and Wells 2002). But, does this imply that humans' digestion is an extended process just because some of the functions originally performed by our ancestors' primitive digestion apparatus are now externally carried out in the cooking process? Sterelny claims that such external process is not a *constituent of* but a *complement to* our internal digestion.

Following this analogy he argues that, just as our digestion is complemented by external technologies and similar structures, also our mind is complemented by many external features we find or build in our niche and that implement our cognitive processes. This makes Sterelny tend to converge on the second wave positions I analysed in the first chapter (see *supra*, I, § 2); he especially feels at ease with John Sutton's perspective about distributed cognition and the use of environmental structures as a support for cognitive processes (Sutton 2006a, 2010). For this reason, Sterelny prefers to endorse a *scaffolded* conception of the mind instead of an extended one.

The interesting point of applying this framework to the discussion about the mind I have been developing so far is that the Niche Construction Theory introduces the *abiotic elements* of the physical space an agent inhabits in the evolutive dynamics which characterize her. In § 2 of this chapter I highlighted niche construction contribution to the explanation of the feedback loop between cultural and genetic inheritance: an agent can codirect her own evolution developing cultural responses which eventually affect her genes through the mediation of an abiotic component, an ecological factor. In the same way, a concrete ontogenetic information can affect an heir-generation through the mediation of abiotic factors which are present in the niche.

Sterelny's contribution to the debate on extended mind, then, is that our cognitive practices are after all one of our specific ecological engineering strategies. The clearest example to illustrate this point is the cited case of *apprenticeship*: when an apprentice joins the workshop of a woodcrafter, she is oriented by the expert artisan in an adequate learning space; she is told where tools are and which ones are the most appropriate to the work she has to carry out. In the same sense, we are born in an environment already scaffolded by our (genetic or non-genetic) ancestors: we are born in an environment which has already been engineered in such or such way and we take advantage of the ecological, social and cultural inheritance available into our niche.

Considering this context, my point is remarkably different: Sterelny sympathizes with the second wave focus, namely the complementarity of internal cognitive processes with external resources working as scaffolds for cognition; nevertheless, I believe that niche construction framework strongly supports the third wave claim that cognition is mainly a collective endeavour. In effect, even if Sterelny defends a social dimension of cognition as well, he clearly sees a difficulty for the extended mind in the fact that our external resources are not necessarily individualised (Sterelny 2010, p. 476), which in his perspective implies that these resources cannot be considered as *constituents* of our minds. In the same way, even though we can use our fellows as sources of crucial information, in his opinion we have no reason to consider them as an extension of our mind. He admits that cognition in humans is generally a social (or at least a socially-supported) activity, but he tends to discard the possibility of a collective agent where a group of agents use and share the same external resources (Sterelny 2010, p. 477, especially footnote 7). This is why, I suppose, he does not commit to the idea that a shared mental structure emerges among the members of a community.

I do not think that Sterelny's point is a priori incompatible with my own; actually, I do endorse the scaffolded view he proposes. However, I see just there a good ground to set up a proposal in defence of collective mind: as I argued all along the last chapter, we have a collective mind when collective mental properties emerge in a human group; in this case, such collective mental properties are collective memories and beliefs and, eventually, the very semantics which enables individual agents referential

practices, that is, their meaning-making skills. These emergent properties are implemented by publicly shared representations emergent into the collective. Such representations are what ultimately constitute the cultural niche.

Sterelny also notes that Clark and Chalmers, in their famous example of Otto's memory problems, commit to the parity principle (see *supra*, I, § 1.1) and are then forced to assume a strict, individualized and automatically endorsed relationship between Otto and his notebook in order to prevent external interferences in Otto's extended mind. In fact, if other people could access Otto's special mental device, he could never be sure of his own 'beliefs' which would possibly be subject to an alien deceptive intervention. But it appears as if Clark and Chalmers do not consider, in their original example, that in many cases of our daily experience we find that external information is trustworthy just because it is shared by many people in a public space:

> «[I]nformational resources in a shared space are sometimes reliable *because* they are shared. Such resources are the *joint product of many agents* [my emphasis] and are typically used at unpredictable times and places. In many circumstances, public domain resources cannot safely be used to manipulate a specific target for a specific purpose. While it might be possible to derail Otto's social plans by erasing a crucial line in his notebook, it would be much harder to do so by, say, corrupting the tools he uses to negotiate the subway system. Maps of the underground exist in many copies, with many users able to spot fakes. It is a dispersed, highly redundant resource» (Sterelny 2010, p. 474).

It is exactly this one the crucial point: external resources are mostly *redundant*. This makes the information they carry widespread and persistent. Like an ecological inheritance.

Now, ecological inheritance may be physical when it concerns ecological variable or developmental phylogenetic artefacts (e.g. beaver dams or bird nests); instead, when it refers to the social or cultural niche, it is strictly dependent on the community of agents who constitute the social and the cultural niche. This is particularly evident in the case of the cultural niche, because its components are made up of *semantic relations*,

such as the material symbols Sterelny appeals to, which ultimately consist in that meaning-making activity that – following Gallagher's interpretation – I intend as cognition.

So, my point here is that the emergent cultural niche a human community gradually builds is the space where collective memories and collective beliefs, as I connoted them in § 3.1 of the last chapter, take place. They are public shared representations available in the niche as a kind of redundant ecological inheritance, which eventually affects the behaviour of the other dwellers of the niche.

Then, differently from Sterelny's focus, I do not see the ecological space a human community inhabits just as a place where many agents collectively develop cooperative strategies and cultural devices which *complement* their individual capacities. Instead I remark that these semantic relations, which enable the very existence of a cultural niche, eventually constitute the collective mind of the human community where they emerged, a system whose individual integrants are involved in cognition in so far they implement these emergent meaning-making processes.

Therefore, I claim that the meaning-making activity, which is a peculiar feature of human cognition, occurs into a *structured space* like the ecological (and social and cultural) space of the apprenticeship; the 'scaffolds' of such structured space are those publicly shared *representations*, where meaning crystallize, and the *canonical codes* (see *infra*, V, § 3) which regulate their content. Consequently, I claim that publicly shared representations, displayed by means of representational artefacts (see *infra*, IV, § 1), ultimately allow a human community to accumulate a canonically codified *cognitive capital*, which exerts a synchronic influence into the niche among the individuals of the same generation and a diachronic influence on the future generations, being these representations instantiated by concrete artefacts a real *ecological inheritance*.

I my account, then, meaning production turns out to be a collective endeavour complemented by the use of public representations, which ultimately enables the emergence of collective memories and collective beliefs (see *infra*, V, § 3).

In the next chapter I shall develop further this theoretical framework with a special focus on artefacts and collective memory.

CHAPTER IV

Artefacts, exograms and the embracing memory

In the last chapter I presented in detail the Theory of Niche Construction: I remarked the novelty of this approach for what concerns the evolution of species, namely the introduction of a coevolutive loop between a certain agent and her concrete environment. Also, mutual relations among the agents operating in a determinate niche (and their consequent evolutive outcome) represent a radically novel feature introduced into the debate by niche-construction theorists.

As I highlighted along the last chapter, the niche appears to be the physical space an agent builds all around her when trying to improve her fitness, triggering processes which can affect both in a positive or negative way that very organism. This framework entails, then, that also phenotypes matter in the evolutive process: in fact, information gaining and transmission is not a merely genetic fact. That is, phylogenetic learning is relevant for evolution, but ontogenetic learning is relevant too. Effectively, idiosyncratic adaptive strategies result from ontogenetic learning (like in the case of Galápagos finches treated before), but it is even more relevant the fact that such ontogenetic information has a concrete effect on the niche and, through this redesign of the niche, this information can be transmitted to the following generations. This is what niche-construction theorists have defined as *ecological inheritance*.

Ecological inheritance really is the strongest nexus connecting agents within the niche and providing a real support for my theory of the emergent collective mind of a community, the cultural structure of external information which takes place where the group builds its niche. As I explained in the last chapter (see *supra*, III, § 2), ecological inheritance does not necessarily consist in physical elements, like in the case of the beaver's dam; rather, it also includes all the other agents sharing the same physical space and all their *actions*. In this sense, when intersubjective relations become stable into one's own environment, these social elements also represent a form of ecological inheritance for the following generations.

Finally, when a cultural niche emerges, the ecological inheritance starts to include *material symbols* too and, as Sterelny notes, this might indicate the advent of the symbolic thought as well, which is the mark of the behavioural modernity (see *supra*, III, § 2.1 and § 3). Material symbols, by the way, are *artefacts*; this means that they are physical elements deliberately produced by an agent when modifying her niche, as I noted in the previous chapter. But the term 'artefact', in ecological niche construction, refers to a wide category of things, including nests and burrows, dams and traps. Instead, in a cultural niche, artefacts are generally *tools* (see *supra*, III, § 1.2) and the word 'artefact' includes also all those things which, like material symbols, feature representational attributes. This class of artefacts has been defined as *cognitive artefacts*.

In this chapter I shall then argue that artefacts carry out a crucial cognitive role in human dynamics, being a material support for external records of information. This enables artefacts to display public representations into the community's niche. As I previously argued (see *supra*, II, § 1.2.2), collectively built and consumed representations are a main feature of a collective mind; representations are publicly shared in a physical space through the concrete artefacts they crystallize into; sharing such public representations by means of artefacts enables the emergence of a collective mind's specific features, namely collective memories and collective beliefs.

After providing a definition of cognitive artefacts in the next section, I shall compare it with the notion of 'exogram', coined by Merlin Donald (1991); I shall argue that an exogram is just a particular case of cognitive artefact.

Subsequently, I shall develop the notion of exogram in terms of an artefactual support for collectively shared external information and, then, for collective memory and beliefs.

A final section will work as a connection between collective memory and beliefs on one hand and, on the other hand, *public narratives* affording a normative structure to the agents living in the particular cultural niche where both collective memory and beliefs emerge from the community.

1. Cognitive artefacts: definition and theoretical reach

In his seminal work *Cognitive artifacts* Donald Norman introduced the first definition of what a cognitive artefact is: «A cognitive artifact is an artificial device designed to maintain, display, or operate upon information in order to serve a representational function» (Norman 1991, p. 17). This definition is important for my objectives because, in Norman's opinion, what ultimately defines an artefact as cognitively relevant is its capacity for displaying and manipulating representations[1] which, in my proposal, enable meaning-making activity.

Then, he considers that, on one hand, human beings are massive environmental engineers who develop and employ many kinds of artificial devices in such a task; on the other hand, he remarks that we humans feature a crucial skill for transmitting to the next generations all the changes we operate and accumulate in our environment (*idem*). This is, in my opinion, just a different way to say that we are powerful niche-constructors and that we manifest an uncommon ability for the production of a huge, structured ecological inheritance we eventually bequeath to the following human generations who will inhabit our niche in the future. And I would like to stress, here, that cognitive artefacts are very powerful engines for this kind of niche construction, like many other cultural processes (see on this point Odling-Smee, Laland, and Feldman 2003, p. 240).

Artefacts are central in our daily life, but they were also crucial all along our biological evolution (Jeffares 2010). This is because artefacts in general are designed to enhance human biological capacities, dramatically boosting our daily performances. This is true of many kinds of artefacts which increase our speed, make us stronger or expand the varieties of the environmental resources we can take advantage of (Norman 1991, p. 17).

[1] In this section I deliberately take into account Norman's definition of cognitive artefact as the main reference, being it foundational for the subsequent theoretical discussion about cognitive artefacts arisen in the last thirty years (for instance, see Cole 1998; Heersmink 2013, 2017; Omicini, Ricci, and Viroli 2008; Theiner, Allen, and Goldstone 2010; Theiner 2011). Also Heersmink (2013, pp. 471-472), reporting three more definitions of cognitive artefact, remarks the representational criterion established by Norman as determinant.

But, if compared with artefacts in general, cognitive artefacts are special in the sense that the human capacities they boost are *cognitive capacities*. This is because cognitive artefacts are designed *to manage information* which may be crucial for a human agent: «When the informational and processing structure of the artifact is combined with the task and the informational and processing structure of the human, the result is to expand and enhance cognitive capabilities of the total system of human, task, and artifact» (Norman 1991, p. 19).

So cognitive artefacts, in Norman's words, «expand and enhance» our cognitive skills, but they do not enhance human cognition through a mere amplification of raw biological capabilities; rather, they entail a change in the nature of the task to be accomplished. Norman illustrates the difference between the enhancement afforded by a normal artefact and the enhancement provided by the use of a cognitive artefact through a concrete example (*idem*): a megaphone amplifies the voice of an agent, but it does not affect her cognitive proficiency; instead, the use of a written language or mathematical notations radically influences an agent's cognitive skills redesigning her task-environment (Kirsh 1996, p. 422ff).

Then, I remark that the introduction of a cognitive artefact as a third term into the originally dual relation between an agent and a cognitive task, enables an *epistemic action* (Kirsh and Maglio 1994), and this is pretty clear in the case of mathematical notations whose application implies the reorganization of a determinate cognitive task: for instance, the mental manipulation of quantities (Menary 2007, 2010a, 2010b; Menary and Gillett 2016).

To reconnect this point with my argumentation involving artefacts in previous chapters (especially in chapter II), it is useful to reconsider the case of the check-list/notebook: both Otto's famous notebook and Jane and Jack's grocery list (this last, an example of my own) are artefacts of the check-list kind, as Norman calls it (1991, p. 20), that is a list of items graphically represented on a certain surface, such as a paper sheet. Thus, Norman notes that any system [agent + cognitive artefact + cognitive task] can be analysed both in terms of the «system view» (Norman 1991, pp. 20-21) and in terms of the agent's view to describe the effect the use of the artefact provokes: focusing

on the system, the artefact looks like enhancing cognitive skills making memory more powerful (an agent is now able to remember many more things for longer); rather, considering the agent's perspective, the cognitive task is now radically different, for she does not anymore 'remember' the items recorded in the list, while she just has to 'read' and 'check' the items in her list. This is to make clear that the involvement of an artefact into a cognitive task implies a *qualitative change*, not a quantitative one.

Then, the main feature which makes an artefact serve a cognitive function, in Norman's account, is its capacity to implement representations. It is possible to distinguish cognitive artefacts in terms of their representation displaying mechanisms, so we have *surface representation* and *internal representation* artefacts (Norman 1991, p. 25).

Internal representation artefacts store their informational content in some internal part of the device, while they display such information on an appropriate surface when it is requested for an agent to use it: we have an illustrative example in a computer, an artefact which stores data in its internal memory and circuits, then it displays this information onto a screen, so that an agent can read and use it.

Surface representation occurs, instead, when the very visible surface of a certain artefact is the support onto which representations are displayed: a very familiar example is the paper sheet an agent writes onto; in fact, graphical signs are traced, stored and manipulated on its surface. Clearly, both a map and a book are surface representation artefacts, because they display graphical information on their own surface. And, I say, also a painting should be considered as a surface representation artefact, considering that it stores and display its informational content (in this case, a visual information) on its own surface (see *infra*, IV, § 4). This last kind of artefact will appear crucial for my argument when I shall analyse, by means of a case study, the normative role of public representations shared by a community[2].

Consequently, it is important to remark that Norman also introduced, in his seminal work, the concept of *representational naturalness*, to mean the appropriateness of a surface representation for referring to its object (Norman 1991, pp. 28ff). This is to say

[2] This point will be extensively treated in chapter VI.

that representations do not work arbitrarily: for instance a map, to effectively illustrate the distribution of population in a certain area in terms of density, has to reproduce such additive relation through a graphical element, such as the increasing density of colour hue (*ibidem*, pp. 30-31): a map which used a denser colour hue to represent a low concentration would likely fail in its purpose.

Thus, I think that representational naturalness does not seem so different from the classic semiotic concept of *iconicity*, which refers to the grade of isomorphism existing between a certain representation and the object it refers to[3]. This point is particularly relevant for my argument that public representations enable meaning-making, because Norman considers that a necessary connection exists between the *structure* of a representation and its *referent*, at least when the representation is successful.

In fact:

> «The form of representation used by an artifact carries great weight in determining its functionality and utility. The choice of representation is not arbitrary: each particular representation provides a set of constraints and intrinsic and extrinsic properties. Each representation emphasises some mappings at the expense of others, makes some explicit and visible, whereas others are neglected, and the physical form suggests and reminds the person of the set of possible operations. Appropriate use of intrinsic properties can constrain behavior in desirable or undesirable ways» (Norman 1991, p. 34).

The form, in Norman's words, determines the functionality of a determinate representation. This is because each representation features *particular properties* which constrain the agent's actions: a representation can emphasize some elements instead of others, indicating a particular use as more appropriate than another one. The physical form ultimately *suggests* possible operations to the agent.

Norman calls the intrinsic properties of a representation conditioning agent's behaviour *forcing functions*. These seem almost equivalent to Gibson's concept of

[3] A wider comment about this point will be offered in chapter VI.

affordances (Gibson 1979), but Norman's terminology stresses that they *constrain* the action, they do not simply offer possibilities[4]. This point is important for it will allow me to argue that the meaning of a representation is strictly dependent on its very structure: that is, *the meaning is constrained by the form* of the representation (or by its *shape*, when it is an image), so that the meaning is grounded on intrinsic, iconic properties of the representation (see *infra*, VI, § 4.1). Hence, representations and objects are not so different after all: «[a]ny design can be thought of as a representation», says Norman (1991, p. 34). Both of them manifest forcing functions which limit the possible intentions of an agent using them.

The fact that the form and the meaning of a representation are interdependent is even clearer, in my opinion, when we consider that a representation can eventually express a meaning its author did not mean to; in fact, «[r]epresentations carry with them many subtle intrinsic properties, often ones not intended by the designer. Line lengths represent quantity, and two lines of different lengths thereby intrinsically present a comparison of the lengths, even if that is not intended by the designer» (*idem*).

Thus, as the meaning of a representation is constrained by its *form* (namely, its intrinsic forcing functions which constrain its appropriateness), so the meaning of an artefact is constrained by its proper *function* (namely, how it is used or how it is usable). This same conviction encouraged the philosopher Richard Heersmink a few years ago in the attempt to sketch an artefactual taxonomy, in order to provide criteria to identify cognitive artefacts; then the function each of them carries out turned out to be the fundamental criterion for the organization of such a taxonomy (Heersmink 2013)[5].

[4] About the difference between affordances and constraints in Norman's work, see (Norman 1999 and 2013, chapter 4). Also (see *infra*, V, § 1).

[5] It has been criticized the necessity and even the mere possibility of a 'category' of the artefacts, meant as an epistemic kind (Vaccari 2017). I do not pretend here to go deeper into this interesting discussion, because it exceeds the limits of the object of my thesis. I am using here Heersmink's taxonomy because it is functional to illustrate some crucial qualities of representational artefacts in particular, which eventually are central for my thesis about the collective mind, being them a main support for information shared within a cultural niche and implementing the emergent mental properties of a human community.

In a situated cognition perspective artefacts play a determinant role and mark our way to be cognitive agents. Thus, Heersmink noted that the very function an artefact carries out as a complement to agents' cognitive activity is what ultimately defines it.

Situated cognitive systems have interactive and integrated components which form together a wider system (Sutton 2006a). They can involve artefacts or not: for instance, transactive memory systems are situated cognitive systems which supervene onto dyads of individuals or groups (Theiner, Allen, and Goldstone 2010); nonetheless, different situated cognitive systems consist of human agents and artefacts.

First of all, artefacts represent just a sufficient but not necessary condition for a system to feature situated cognition (Heersmink 2013, p. 467). In fact, as Edwin Hutchins reminds us (1995, 2011, 2014), we humans take advantage of many resources for our cognitive necessities: for instance, internal (or internalized) cognitive structures and external scaffolds like in his famous example of star-guided navigation[6].

There is certainly a notable difference between cognitive artefacts and cognitive external resources in general: effectively, cognitive artefacts are after all cognitive *technologies* (namely, physical objects an agent can use), while the resources Hutchins mainly points out are basically cognitive *techniques* (that is, methods or procedures to get a cognitive purpose). They are both artificial but only technologies are artefactual, then manually designed physical objects (Heersmink 2013, p. 468); while natural resources used as artefacts have been defined as *naturefacts* (*idem*): this is for instance the case of constellations used as a nautical map.

In this framework, both artefact and naturefacts appear as playing a relevant role in cognition; in particular, artefacts played indeed a main role in our evolution, being much more malleable in information storing and manipulation. What makes them special is that their properties can be *redesigned* to serve better their scope. So, drawing on Norman, even if each artefact carries intrinsic not intended properties, I remark that it can be restructured by its author in a way that *selects the properties* the agent needs.

[6] This consists in the description of Polynesian star-guided navigation: a Polynesian mariner uses constellations as points of reference to trace her route.

These artefactual properties are so relevant just because they do not replicate the agent's biological cognitive properties; they rather integrate the agent's normal cognition. And this integration is fruitful just because of the different properties involved in cognition: for instance, exograms are relevant because they support and complete engrams[7] (Heersmink 2013, p. 470, drawing on Donald 1991).

Moreover, while it is true that many artefacts are not primarily designed to serve a cognitive scope such as a hammer, a pan or a spoon, it is indeed true that any artefact *can acquire* a cognitive value depending on the (even contingent) *function* it is carrying out (*ibidem*, p. 471): a carpenter, for instance, can put on her worktable an empty screw packet to remember she needs to buy some more screws; so the screw packet can acquire a mnemonic function, like a knot in a handkerchief[8].

So Heersmink notes (and this is novel with respect to Norman's definition) that cognitive artefacts can manifest both a representational or a non-representational nature, but in any case they have to feature an *informational structure* (*ibidem*, p. 472).

To make clear what 'representational' means in relation to artefacts, he endorses the well-known Peircean triadic scheme where signs (here to be understood as representations) are recognised as iconic, indexical or symbolic. Some tokens which may clarify this definition would then be a map for the species of iconic representational artefacts, because of its isomorphic relationship with its referent; then a compass would fairly represent the indexical species, because the information it encapsulates consists in an arrow pointing at a certain direction; finally an abacus works as a good token for the symbolic species, because the informational structure it contains maintains just a conventional relationship with the referent (Heersmink 2013, p. 474).

Non-representational artefacts, instead, encapsulate *ecological information* which can be spatial or structural. An artefact affords non-representational spatial information to an agent depending on its relative location in the environment: for instance, the empty screw packet I told about before works as a mnemonic cue for the carpenter through the spatial information it affords to her. Instead, a non-representational

[7] Respectively, external and internal records of memory (see *infra*, IV, § 2).
[8] It is appropriate to remember here, about this point, Vygotsky's fragment I quoted in chapter II, § 3.1.1.

artefact can afford structural information when its structure is manipulable. Heersmink, drawing on Kirsh, uses the example of Scrabble tiles: tiles can be rearranged in different sequences to prompt word recall and, depending on the sequences which are formed at any moment, new information emerges (Heersmink 2013, p. 479).

So, representational cognitive artefacts are involved in a triadic relationship with the world, for they mediate the cognitive relations an interpreter maintains with the world so that [interpreter ↔ artefact ↔ world], while non-representational artefacts implement dyadic relations like [agent ↔ artefact] (*ibidem*, p. 477). Eventually, structural information recalls Norman's point I endorsed and drew on before, namely that the structure of an artefact contains intrinsic properties which carry information not intended by the designer. I said, about this point, that ultimately the form of a representation defines its meaning; so, a structural change into the form of a representation would entail a change in its meaning too, just as it occurs in the case of the rearrangement of letter tiles in Scrabble: changing the position of even one tile could imply a very different meaning of the whole sequence. The same happens with iconic representations, as I shall explain in the last chapter.

Having now a clear idea of what a representational cognitive artefact is, I shall focus in the next section on a particular kind of such artefacts: namely, exograms.

2. Exograms: a peculiar kind of cognitive artefact

Among representational artefacts there is a particular typology which appears extremely interesting for the aim of my research: these artefacts are *exograms*.

'Exogram' is a neologism introduced in the literature about cognition by the Canadian psychologist Merlin Donald, almost thirty years ago (Donald 1991). It has been coined as a symmetrical concept with respect to the one of 'engram', this last coined by the American neuropsychologist Karl Lashley in the 1950s to refer to internal records of memory (Donald 2010, p. 71). Therefore, «[m]emory records stored outside the nervous system (for example, clay tablets, papyri, printed books, government

archives or electronic data banks) can be called 'exograms'» (*idem*) and they can be very different in format and kind.

Donald's work inserts itself in an evolutive framework about cognition. He studied the evolution of human representational skills and offered a focus on three different chronological phases in the evolution of human representational mechanisms. I shall not discuss here Donald's arguments in favour of such chronology, nor I shall commit in any moment to the modular approach he has about cognition, drawing on Fodor's arguments. My focus, here, is about the use of exograms in the collective construction of shared knowledge in humans.

So Donald's point is that, during our evolutive process, we humans developed several representational mechanisms to deal with the challenges of daily life[9]: firstly, we developed *mimesis*, which is meant as the ability for representing and communicating a semantically relevant content by means of one's own body. Gestures imitate what someone wants to represent (e.g. a movement) and body sounds echo those of the referent (e.g. an animal call). Hence mimesis looks like an analogical strategy for representation which refers through perceptual similarity, by means of iconic and metaphorical gestures (Donald 1997). Mimesis, in Donald's perspective, is the first stage of the human attempt to transcend the limits of «episodic» communication (Donald 1998a, p. 14). It represents the basis for rhythmic representations such as dances and pantomimes.

More recently in evolution, Donald places the development of structured verbal representations, which afforded a basis for the development of the narrative thought: that is an oral public transposition of reality, which worked both as a recall tool for public memory and as a regulative tool, providing standardized representations of the world. This ability represented in Donald's account a very important change in human

[9] Here I am going to summarize in a very short sketch Donald's chronology of the evolution of human representational capacities. This does not mean that I commit to his account of such an evolution: in fact, I do not have any pretension, here, to explain the phases of human representational development or the like. Instead, I am reconstructing here Donald's chronology just because it makes sense of his idea of exograms. The concept of exogram, in fact, is what really matters for my account. Then, I am using Donald's chronology as a mere framework to explain the role he attributed to exograms.

management of information and, ultimately, it allowed the emergence of mythical narratives. These last were based on biological brain devices for mnemonic recall (*idem*).

Finally, during the Upper Palaeolithic, humans started to develop external storage mechanisms (Donald 1998a, p. 15), adopting new costumes such as the use of material symbols. And this is coherent with what I have reported before from Sterelny: namely, the relatively late development of a symbolic culture in humans (see *supra*, III, § 2.1), which supervened on the wide use of material symbols[10].

Then exograms generally are material supports for information which exist in an external physical space, namely the ecological niche a human group builds all around it. Subsequently, the existence of material symbols enables the development of symbolic codes and the supervenience of cultural structures which eventually constitute the cultural niche of a human community. Written systems, for instance, are this kind of symbolic code (Harris 1989; Menary 2007; Donald 2010). So, the development of simple symbolic systems at first and the posterior advent of literacy enabled humans for the externalization of biological memories into physical records, improving the information-storing skills of the humankind.

Exograms, by the way, are relevant for human cognition mainly because they feature very different properties with respect to engrams: in fact, while engrams work through physiological systems which are fixed, exograms are instead unlimited in their physical support typologies and mechanisms; engrams show a constrained format, while this is not the case of exograms, whose formats are almost unlimited; while engrams manifest a limited capacity for memory storing, exograms manifest an unlimited storing

[10] Someone might object that a symbolic culture could have been enabled by the previous development of language in humans and, consequently, this would have allowed the posterior emergence of a material culture. Nevertheless, saying 'symbolic culture' is not the same of saying 'abstract thought': in effect, linguistic skills might have enabled the abstract thought, which successively permitted the development of material symbols, but 'symbolic culture' rather appeals to the emergence of symbolic *codes*.
Consequently, it seems more plausible, in my opinion, that material culture developed in human niche as a side effect of some kind of previous material structure whose function was not symbolic at first. For instance, let's consider once again Sterelny's example (see *supra*, III, § 2.1): some material element of the human niche, like body-painting, would likely have been used as a mimetic technique in hunts at first; but painting traces on bodies afford (see *infra*, V, § 1) 'marking', so they could have implemented a symbolic marking system for group membership. So, material structures in human ecological niche could have supported the emergence of a specific symbolic function from an early associative relation, which lately could have developed in a real instance of material culture (e.g. a specific set of marks with a specific meaning or a particular body-painting style).

capacity; retrieval is weak in engrams, while it is highly reliable in exograms. Finally, the information stored in an exogram is highly manipulable, this is not the case of engrams, whose biological storing mechanism impedes a direct access and intentional manipulation of their content (Donald 2010, p. 72)[11].

But, what is mostly relevant for my own approach is that the advent of exograms liberalized information, making possible a wide diffusion of ideas collectively shared by many individual brains. So the use of exograms allowed the redistribution of cognitive work in society, dramatically boosting memory capacity or *redesigning cognitive* tasks, like in the case of mathematical notations: in fact, «the best exographic systems reduce the load on the brain by simplifying some operations, and designing the interface technology so as to focus the mind on task-relevant issues. The juxtaposition of mind and exogram quite literally changes the nature of the task facing the brain» (Donald 2010, p. 76). However, Donald remarks the centrality of the human nervous system as the ultimate driver of the exographic system (*idem*). In fact, in his account culture appears to be after all «a gigantic search engine» (*ibidem*, p. 78), which organises the collective knowledge of individuals in human communities.

Therefore, in Donald's view, the advent and diffusion of exograms entailed a deep change in the most intimate way humans store and manage the information they recollect: while biological memories, once language appeared as a stable phylogenetic trait in humans, were registered and organized through oral mnemonic techniques (Donald 2007, p. 220)[12], the introduction of exograms decoupled, in some concrete sense,

---

[11] The notion of engram, here, has to be considered as an operational term: originally, it was meant by Lashley (1950) as a memory record stored somewhere in the nervous system; however, such a 'unit of memory' as never been identified. Donald uses it to refer to brain-based memories in general, and he casts the concept of exogram as a merely symmetrical notion, namely, an external artefact-based record of memory.
Perhaps someone sympathetic with the claim of the 'mark of the cognitive' (see *supra*, I, § 1.1) might wonder if an engram is needed to interpret the content of an exogram, but Donald simply did not focus on this issue: an exogram simply is a codified piece of information, encrypted into some external device, that a human agent can decode. Its content (its meaning) may derive from the *function* a group of users assign to it and this is, actually, the view I endorse.

[12] Certainly, oral mnemonic techniques are not exograms. They are strategies for memorization which are not grounded in material supports: for instance, those structural properties of poetry or 'oral literature' (Havelock 1986) such as rhetorical figures, rhymes and metrics. By contrast, exograms are always artefact-based memory records.

that information from any particular brain[13]. In Donald's account human cultures appear to be *cognitive infrastructures* which connect many individual brains in the same information net: they «interlock many minds into social meta-organisms [which] function as 'distributed' cognitive networks» (Donald 2007, p. 215). Nevertheless, this does not imply for Donald that human cultures constitute instances of collective minds, because they do not feature nothing like a self, a personhood or a phenomenological consciousness (*ibidem*, p. 219). More specifically, he says:

> «Institutions and social organizations are not conscious entities and we cannot say they have minds. But they are cognitive entities and they do perform cognitive work. They have beliefs, perceptions and plans. They evaluate situations, and react creatively to challenges. Although they cannot function without the individuals that make them up, institutional structures rarely depend on single individuals over the long run. They dominate the minds of their members, and individuals assimilate institutional values to such an extent that they rarely violate them» (Donald 2007, pp. 220-221).

Donald does not acknowledge the status of 'mind' to a collective cultural structure because his criteria for determining what a mind is are based on what is meant to be an individual mind. But, as I have been arguing so far, collective minds do not replicate individual minds features. In particular there are no theorists, not even among those sympathetic with a collective approach to cognition, who would endorse a phenomenological account of an alleged collective consciousness (List 2015).

---

[13] Andy Clark argued in favour of language as «the ultimate artefact» (Clark 1997, chapter 10), considering that we use language not merely for communication, but more importantly «as a tool that alters the nature of the computational tasks involved in various kinds of problem solving» (*ibidem*, p. 193). Language is then conceived, in his account, as an external device. In this sense someone might observe that language itself, inasmuch as it is conceived as an artefact, is what firstly decoupled information from concrete subjects. But I disagree on this point: in fact, even though language can be decoupled by this or that concrete speaker, it is never really decoupled by a concrete *community* of speakers (and Clark himself maintains a brain-centred view of such an artefact). The only way in which information (including *linguistic* information) can persist as decoupled by any concrete user is when it is externally stored into a material symbol, an artefact which becomes, for that reason, an exogram.
Indeed, I take communication for granted into the considered group. I am not contending, here, that exograms *constitute* communication; I just claim that even a piece of linguistic information needs to be encoded into a concrete material symbol to persist as decoupled from concrete users: for instance, a *written* sentence, which certainly is an exogram.

What instead supports my thesis, in Donald's view, is also the public role he acknowledges to mythological narratives, namely «system[s] of governing myths, which regulated every aspect of daily life» (Donald 2007, p. 220). He recognises that public narratives «dominate the minds» of the integrants of a community in their *daily life*. This is because «[t]he culture establishes the environment within which the ontogenesis [takes] place» (Donald 1998a, p. 11) and individuals contribute to the construction of such environment: they modify with their particular representations their cultural niche, then such modifications feed back to individuals, changing their particular representations. For this reason I shall argue throughout chapter VI that public narratives such as myths do not just regulate practical life of people; they also regulate the way people form categories and references. In this dynamics, collective representations exist as *decoupled* from any concrete agent just because they are carried by exograms, namely, physical artefacts a certain information is encapsulated into.

3. From external records of information to a collective approach

The 'exographic' revolution (Donald 2010) afforded humans the possibility to externalise their memory and easily share information in groups through the use of artefacts in cognitive operations; cognition became even more a collective fact based on material culture, and networks of exograms and individuals formed concrete instances of distributed cognitive systems (Donald 2007, p. 219ff; Sutton et al. 2010, p. 525). Exograms, being a support for the external recording of information, represent the physical basis of the distributed cognitive network whose nodes are the individual agents of a certain community (Donald 2017, p. 205). Exograms, eventually, work as cognitive artefacts because they make part of a complex web: the information they carry makes sense because each exogram is after all «a node in a dynamic social-cognitive system. [And] The system itself defines the role of such objects [in the web]» (Donald 1998b, p. 185). They carry out their cognitive function because they exist in a *social space*.

As I remarked before (see *supra*, I, § 3), in the third-wave approach to extended mind the 'mental' is a category with a strong social dimension. A Vygotskian radically

social account of the mind (see *supra*, II, § 3.1.1) is therefore a common background for theorists sympathetic with the Third Wave claims. This is true for Donald as well (Donald 2000, 2001, 2007, *passim*).

Now, it is important to remark, here, that Vygotsky's focus was not on collective minds and the like; he never wrote about minds as emergent features of groups. His focus, in fact, was more on the individual and the socio-cultural influence which affects the development of her mind. Therefore, his perspective on this issue was from the outside to the internal cognitive processes of an individual mind. For instance, Vygotsky's approach to the analysis of language considered that the child's development was mediated by the internalization of linguistic structures got from the external social environment: after a first step in which the child organises her activities through confabulation (egocentric speech), she internalises that in a private language (internal speech). So, in his account, the internal speech serves a *self-regulative function* and the child gets it from the social environment she grows up within. Then, he stressed that the individual mind is *culturally entrenched*.

The third-wave postures, instead, draw on this concept of a culturally entrenched mind to argue that environmental socio-cultural factors ultimately constitute mental processes. I, myself, stress this interpretation from Vygotsky's thought.

The Vygotskian idea that the mind was culturally entrenched, radically opposed Jean Piaget's idea of the 'isolated' mind (Vygotsky 1934/2019, chapter 2). For the Russian psychologist understanding is the result of a social activity, it is not an individual invention (Cole and Wertsch 1996, p. 250). In this social activity, artefacts carry out a crucial role (*idem*): in fact, the actions of individual agents are *mediated* by artefacts and such forms of mediation are transmitted to the future generations through *cultural structures* (*ibidem*, p. 252). For Vygotsky both the environment and the agent maintain an *active* reciprocal relationship: the child, during her development, receives many inputs from her physical and social environment, and reacts and changes such environments with her responses. So cognition is possible in this framework because of such dynamics of *co-construction* which involves a third determinant factor: the products accumulated in the environment by the precedent generations and which constitute the

culture of the human community a certain agent lives within (*ibidem*, p. 251). Hence, a concrete culture constitutes the *medium* through which a certain agent interacts with her environment and, for the same reason, artefacts carry out a crucial mediation role in her cognition (Vygotsky 1934/2019, chapter 5). Actually, it could be said that artefacts do not only help mental processes; they literally *shape* them (Cole and Wertsch 1996, p. 252). They can shape mental processes because artefactual properties are not irrelevant at all for the cognitive task: as Norman argues, in effect, any artefact manifests intrinsic properties which in some sense constrain the behaviour of a potential agent (see *supra*, IV, § 1). Vygotsky noticed this particular feature of artefacts as well. In fact, he reports on a child playing with a stick which he pretended being a horse. Analysing this case the psychologist remarks that it is not true that «the properties of things as such have no *meaning* [my emphasis]. Any stick can be a horse, but, for instance, a postcard cannot be a horse for a child» (Vygotsky 1978/2018, p. 144). The stick, in fact, has some properties that make it *rideable*, some properties that the postcard lacks[14]. Such properties constrain and define the *meaning* of the considered artefact as a horse.

Moreover, in Vygotsky's account of the mind, any cognitive activity is socio-culturally *situated*: there is never a unique tool which is appropriate for all tasks, since 'action' and 'context' are always interdependent (Cole and Wertsch 1996, p. 253). And other people are crucial constituents of our «proximal space» (Vygotsky 1978/2018, p. 124ff), that 'social niche' which offers us a cognitively affordable environment for our problem-solving activity, which generally our educators and caregivers build all around us. Interpersonal processes are therefore a precondition for the development of individual minds, considering that any agent acts into an environment inhabited by other people and saturated by artefacts (Cole and Wertsch 1996, p. 254). I certainly endorse this view and I would like to stress that the collective manipulation of public representations eventually modify even the very categories applied by the individual members of a community (see *infra*, VI, § 4.3).

[14] About the example of the stick/horse, see also Umberto Eco (1975, p. 327). The stick is, in some respects, *iconic* with a horse (or an 'appropriate' representation of it, echoing Norman 1991): it is *rideable* because of *some* of its material properties, such as the length and the rigidness. I shall say more on this point in chapter VI.

This Vygotskian framework looks so close to the Niche Construction Theory approach that it is sufficient to substitute just the terms 'environment' and 'ancestors products', respectively with 'niche' and 'ecological inheritance', to make such similarity explicit[15]: the agent and the environment maintain a co-construction dynamics, and the cultural structure (including the artefacts, like exograms) represents the ecological inheritance bequeathed by the ancestors to the heir-generation. This is like saying that the information contained in the niche shapes the behaviour of the agents which inhabit that niche (see *supra*, III, § 1.2). And it is certainly true in the case of a cultural niche as well: structural information, both in the space and the physical design of artefacts, heads agents' activities (see *supra*, III, § 2; IV, § 1).

In the same line of thought, Vygotsky acknowledges that things *motivate* us, they call for action: «the 'things' dictate [to the agent] what she has to do: a door requires to be opened and closed, on a stair you need to climb up, a bell has to be rang» (Vygotsky 1978/2018, p. 141). In fact «[a]ny perception is a stimulus for the activity» (*idem*).

The ecological niche an agent grows up within is, then, a space rich in stimuli, both structural and spatial information which requests such or such responses from her. Artefacts are a main component of such information which the environment contains and affords. Nonetheless, artefacts implement cognition just because they are involved in a complex network including agents and other artefacts as well (Donald 2007, § 2). As Donald argued (see *supra*, IV, § 2), exograms interconnect many minds in the same distributed cognitive network. In this sense culture appears more as an infrastructure for communication than a subject of cognition; even though, Donald acknowledges that cultures forms, with the agents involved in them, a cognitive entity in a proper sense, a

[15] There is an important distinction which has to be remarked with respect to this comparison: Vygotsky's thought is certainly focused on the *individual* as the basic 'unit of analysis', who is later considered with respect to her socio-cultural context, whose influence is determinant in the constitution of the individual mind. Instead, the Theory of Niche Construction focuses mainly on *groups* of agents and structural properties of the environment to analyse the feedback loops which connect them.
However, Michael Cole and James Wertsch observe that this distinction between 'individual' and 'social' is quite blurry in Vygotsky's account: «In fact, the very boundary between social and individual, a boundary that has defined much of our thinking in psychology, comes into question in Vygotsky's writings. Just as the mind does not stop with the skin in Vygotsky's view, the relation between individual and social environment is much more dynamic than the overly simple division we so often tacitly assume» (Cole and Wertsch 1996, p. 254).

kind of meta-organism made up of individual subjects' minds. He does not talk of cultures as collective minds because he still conserves, as I noticed, a Cartesian bias about the mind, which makes him think of a mind in terms of an alleged 'self' and 'consciousness', so as something featuring a particular phenomenology typical of individual human brains. Instead, coherently with the interpretation of the Vygotskian approach to the social dimension of the mind I have endorsed in the last pages, I shall go a step further to treat the mind itself in terms of a social structure. This is why I focused on peculiar cognitively-relevant emergent properties of collectives, such as collective memory and collective beliefs (see *supra*, II, § 3), that I claim being characteristic traits of culture as a collective mind.

For obvious reasons, then, I endorse the complementarity principle proposed by John Sutton (see *supra*, I, § 2.1): he persuasively argues that cognition is socially distributed (Sutton et al. 2010, p. 524). Artefacts and the environment a cognitive agent operates within clearly form, in his view, real *cognitive ecologies* for the subjects which inhabit them (*ibidem*, p. 526). Sutton and colleagues oppose the 'embedded' account of mind and cognition defended by Rupert (see *supra*, I, § 2.4), which reduces the contribution of external resources in cognition to merely causal inputs stimulating internal cognitive processes (Sutton et al. 2010, p. 532). Instead, their complementarity approach argues for an *interactive coupling* between these two kinds of resources. It is because of this interactive coupling between agents, artefacts and environment that it is possible to talk of cognitive ecologies, applying the model of the analogous case of a biological creature living into an ecological niche.

Sutton and colleagues especially focus on *transactive memory* (Theiner, Allen, and Goldstone 2010) as a representative outcome of a cognitive ecology: they analyse concrete cases where long-lasting couples or small groups of close friends rely on each other to efficiently manage a common store of memories (Sutton et al. 2010, pp. 547-549). Using artefacts and the whole apparatus of their cognitive environment, people making part of a transactive-memory unit share knowledge and gather memories following a clear scheme for division of cognitive labour in the group. This division of

cognitive labour and the awareness of 'who knows what', plus the use of external resources like exograms, enables transactive memory skills in groups of people:

> «A transactive memory system requires both a set of practices and mechanisms of coordination and communication (the process components), and an awareness of the actual or likely distribution of information across individuals within the system (the knowledge component). In ongoing integrative process, the members of a successful transactive memory system will turn what may initially be differentiated knowledge into shared new emergent knowledge» (Sutton et al. 2010, p. 547).

Transactive memory is acknowledged as an emergent properties of *small* groups (Sutton et al. 2010, p. 539; Michaelian and Sutton 2013, p. 7). This is because each member of such «epistemic group agent» (Palermos 2014) is supposed to be aware of the distribution of the different pieces of information managed by some determinate agents in the group. So, it is generally supposed that the integrants of a transactive-memory unit *directly* know each other. For instance, in a small group of close friends who share the same hobby, such as hiking together in diverse natural landscapes, transactive memory could consist in that each of them knows who participated in such or such activity, who was with whom the last time in that peculiar valley on that day etcetera. So each of them knows who might remember what they saw on a certain day in a determinate place. In the same way, in a work-team anyone knows what is the job of each teammate, so it is possible to ask for a report to any specialised operator of the group on the field she is working on.

But external representations allow the extension of transactive memory also to *huge* groups such as wide communities of 'strangers', because 'who knows what' is evident in the explicit manipulation of public representations into an ecological space, and the very scaffolding of such space indicates the distribution of the information within the group. Let's imagine, for instance, a diagram showing the organization of the managers working in a company: the diagram shows the exact distribution of the information within the company to anyone who has access to such public representation.

Therefore, the externalization of memory by means of exograms extend the instances of transactive memory beyond the narrow limits of small groups of close friends. At the same time the use of public representations, as far as it is meant as external recording of information, enables all the members of a cultural community to know what the community itself knows, then what kind of information the community manages and, eventually, how to find and to employ such information: in fact, «[a]ll cultural networks, even those of oral cultures, harness the cognitive resources of many individuals and impose a larger organization, often with a technological dimension, on the mental functioning of individuals. This greatly affects what an individual can [cognitively]achieve» (Donald 2007, p. 220).

Hence, cultural communities are instances of the distributed cognitive ecologies within which we remember (Sutton 2014, p. 410). Remembering is, in fact, a *collaborative activity* (Sutton 2014, p. 422; Williamson and Sutton 2014, p. 113ff), a *situated public practice* which would be impossible in isolation from its environment and context (Sutton 2014, p. 425). Effectively, we remember more by knowing *where* we can find a certain information when we need it than by knowing by heart each piece of information we need (Bietti and Sutton 2015), since remembering is an activity both spatially and temporally situated (Sutton 2014, p. 434). This is because it is enabled by the use of external structures such as exograms or other spatial mnemonic cues and by the support of other people (Sutton 2010, § 3). Cognitive ecologies put the cognitively relevant information all around us. But, when information is publicly shared in a physical space, it becomes verifiable and, then, it acquires a different value as a standard[16]. However, a standard is a *model* and, therefore, it generates a normative reference into the physical environment where it is shared.

This is manifest also in simple cases of transactive memory: in fact, when a particular memory is collectively recalled, piece by piece, by all the people involved in this recalling activity, the collective reconstruction of facts acquires a normative power on the group, regulating the individual memories of each integrant of the transactive-memory unit as well (Sutton et al. 2010, p. 550ff). Moreover, in small groups like families

[16] Information is reliable just because it is shared into a public space (Sterelny 2010, p. 474).

transactive memories can generate and reinforce a collective identity through the autobiographical reconstruction of past events and the discursive negotiation of the correctness of such or such information (Bietti 2010).

I argue that the emergence of normative models is possible because public systems of representations can generate *narratives*; these narratives affect the minds of the individual integrants of a cultural community and they shape, force and head the individual cognitive activity (Donald 2007, pp. 220-221). Moreover, narrative structures are not just made up of mere verbal elements like in oral civilizations; instead, they are made up of visual elements as well: for instances public representations like paintings or frescoes, which indeed show a narrative structure; at the same time, they harness the mind of individual agents by means of their structural constraints and relevant spatial information (see *supra*, IV, §1).

In the next section I shall clarify how publicly shared representations featured by cognitive ecologies create an embracing memory subsisting all around the community, a network of cultural structures which guides the community through its normative power, enabling the emergence of collective memories and beliefs.

4. Public narratives: emergent normative structures in the community

Any community needs a good story to exist. Foundational myths are not mere attempts to explain a mysterious reality whose nature remains mainly obscure for the humans; rather, they serve an identitarian function: they are nuclei of basic 'truths' the members of a certain community automatically endorse and they state *what* the community is and *which place* it occupies in the world. Therefore, foundational myths cast the collective identity of the integrants of a particular cultural community attracting them all through the centripetal force of a strong narrative.

Narratives, effectively, organize collective memories in a coherent structure accessible for all the members of a group. They give to the community from which they have emerged both a spatial and a temporal dimension: in fact, they project such community *somewhere* in the space, *after* a mythological past and *before* a still unknown

future. As Eric Havelock remarked, narratives are the typical dimension of orality and afford to their communities of reference an «oral encyclopaedia» (Havelock 1986, p. 72). Narratives organize the thought of the agents in a certain cultural group, in that network of people and exograms whose cognitive nature was recognised by Donald. They constitute both a *corpus* of collective memories and a system of collective beliefs.

We should not think of a narrative as something immaterial, a monodimensional element whose existence merely develops along the stream of time. Havelock certainly says that oral narratives are «a system based on the echo» (Havelock 1986, p. 84), but, when they are told, they *expand in the space*. When a popular poet used to tell a story in an oral community, it was not a mere 'telling'; it was instead a real *performance*: all around the audience, the bard was *pointing out* the characters she was telling about, like "There was Odysseus, sitting beside Nestor, the wise king" etc.

Appealing to the *mimesis* (Donald 1991, 1993, 1997, 1998a, 2010, 2017) was therefore a constant attitude which expanded the narrative of the bard within the physical space occupied by the audience.

So, narrative information is *transdiegetic* (Floridi 2009, 2011). Namely, it is an information which moves throughout the different levels of a story, within it and *beyond it*, to the audience. It organizes the experience of the community in the space because the narrative thought has indeed a spatial dimension as well: in any story something *occurs* and people *do* things in a certain *place*. Stories always happen *somewhere*.

This is relevant because narratives give a spatial dimension to cognition in general (Havelock 1986, pp. 50-53; Theiner 2011, § 4.4.2). Havelock, for instance, reports on an experiment of Alexander Luria, who discovered that, in an oral culture, the names which were not directly related in a sequence or a list, were memorized in a narrative context as 'actors':

> «When S. was memorizing a series of names he needed a break of few seconds between a name and the following one to have time to *visualize* [my emphasis] the object and to *collocate* [my emphasis] it in a determinate point of a familiar zone, generally on regular intervals along a well-known *path* [my emphasis]. Once this was done, he was able to

walk [my emphasis] along the path starting from its two extremities or any other point and list the objects he collocated on the path» (Havelock 1986, p. 52).

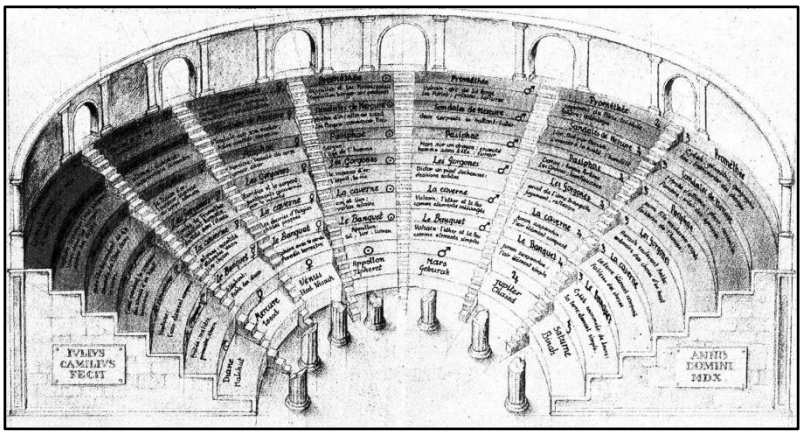

Fig. 4.1 ***The theatre of Memory*****. After Giulio Camillo Delminio**

The mnemonic technique of *loci*, so typical in the Renaissance (Eco 2013), worked just like this. In fact, to memorize a long list of entries (or even an entire discourse) skilled mnemotechnicians used to 'build' in their mind a 'city' or a 'theatre' whose structure worked as the mnemonic scaffold: on the different sectors and steps of the theatre, some content was placed following a precise encoding rule; then, when the mnemotechnician needed that concrete piece of information, she used to 'walk' throughout the structure of the theatre Fig. 4.1 to retrieve that particular content. But this famous image of the 'palaces of memory' dates back to Augustine, at least: in *Confessiones*, X, 8 he depicts himself as entering the *lata praetoria* (huge palaces) of memory, where anything we experience is stored and retrieved when needed, like in an archive or a library (Donald 2001b). Also the image of a 'city' where palaces and streets encoded information were used alike (Eco 2013, p. 8ff)[17].

Hence, these mnemonic techniques are *cognitive techniques* in the same sense introduced by Heersmink (2013, p. 469): they are an internalized equivalent of the constellations used as nautical maps (Hutchins 1995, 2014); they are indeed an instance

[17] Clearly, the contemporary conception of memory stresses that remembering does not consist in merely retrieving some stored piece of information we stock somewhere in the head like a file in an archive: remembering is widely acknowledged, in fact, as a reconstruction process through which we re-build a particular memory from multimodal information. This works also for collective memory inasmuch as it can be conceived as a reconstructive process in which information is reconstructed by means of contextual discursive negotiation (Bietti 2010).
But, when collective memories are attributed to a huge group of people so that they are mediated by the extensive use of exograms, 'reconstruction' has to be understood as the process of following the scaffold structure afforded by a concrete cognitive ecology to retrieve and manipulate the information stored in those external records of information.

of *cognitive ecology* (Huebner 2013; Sutton et al. 2010; Williamson and Sutton 2014; Sutton 2014).

Particularly, Sutton remarks that these «internal cognitive architectures [were] virtual random access systems and internal prostheses which these adepts [the mnemotechnicians] built into their theatres of memory [which] were extended as well as cognitive, even though they didn't happen to be outside the skull and skin» (Sutton 2006, p. 240)[18].

Public narratives in oral cultures show, indeed, this spatial dimension as in the example I quoted from Havelock before. So, for this reason, such narratives should be considered cognitive ecologies as well. In fact, they afford a complex scaffold which complement individual cognitive performances.

Community narratives represent, therefore, a particular case of external collective memory: they enable a community to develop an extended transactive-memory system. Moreover, as I highlighted in the previous section, transactive-memory systems lay the foundations for the emergence of shared representations/memories which work as a standard, a model for the self-regulation of the group: just as a common memory is publicly *built* and *controlled* when it is shared into a small group of close friends[19], also the collective memory afforded by a narrative is publicly built and controlled. In effect, dynamics of co-construction entail also a *co-correction* of the storytelling (Sutton et al. 2010, p. 551).

Finally, narratives are not just orally told through the bard's performance; they can also be 'told' by means of visual representations, iconic artefacts publicly shared within the community, of which there is a plenty of examples in the history of figurative art (Fig. 4.2).

[18] See also (Sutton 2010, § 5).

[19] «In ongoing integrative process, the members of a successful transactive memory system will turn what may initially be differentiated knowledge into shared new emergent knowledge» (Sutton et al. 2010, p. 547). See also (*ibidem*, p. 550ff).

Fig. 4.2 ***Stories of the Virgin Mary*****. After Duccio di Buoninsegna**

I am remarking 'space' as a constitutive feature of narratives, here, because their spatial dimension is what makes them relevant in human cognitive ecologies: it is because they *exist in space* that they constitute an *ecological inheritance*. They are told in a certain ecological niche and they constitute a cognitive scaffold in that ecological niche. This makes narratives important for cultural niche construction, because they are bequeathed to the following generations as performances. I also told about both the spatial dimension internal to a story and the spatial dimension a story occupies in the world, through a performance. This seems not such an automatic connection, at first sight, but my point is that narrations are *naturally spatial* (Havelock 1986, pp. 18, 27, 50-54) because an action always occurs somewhere, so they naturally extend into the space of the audience. *Narration naturally supports performance*. It is never a mere reading; it is always a telling which involves the space the community inhabits, like in a play at the theatre. From a performance to an iconographic cycle representing a story, it is a short step: iconographic cycles are organized as *sequences of frames*, like in Fig. 4.2; they literally are a performance crystallized into a series of juxtaposed exograms shared into a public space relevant for the community.

Just as it happens with an oral narrative, also visual narratives emerge as the outcome of a transactive-memory unit. As I shall argue in the last chapter, in fact, standard representations emerge in a community as models, as *canons*. Eventually, they are the result of the co-construction and co-correction activity of the integrants of the group: a canon is like the 'best version' of a story whose shared experience is collaboratively recalled to memory and co-constructed by a group of close friends. Models are afforded by the representational artefacts that, because of their intrinsic properties (Norman 1991), harness the minds of the individual members of a cultural community (Donald 2007), casting through their canon what I called collective beliefs and supporting collective memories (see *supra*, II, § 3), considering that they are also exograms whose information content has an iconic appearance.

5. Conclusion

In this chapter I have firstly exposed what cognitive artefacts are: I drew on Norman's representational definition of cognitive artefacts. Norman is surely interested in the cognitive role played by an artefact in the economy of individual cognition, but I used his criterion of representationality to argue that some representational artefacts can also support external storing of collective information: in fact, I argued that this kind of artefacts are exograms.

Explaining, then, in the second section the notion of exogram coined by Donald, I used it to underpin my account of collective memories, because I claimed that external records of information can be shared by the different members of the same community. These collective memories can be organized in cognitive scaffolds. So, while Donald speaks of collectives as cognitive entities, denying that they represent instances of collective mind because they do not feature any psychological phenomenology (he has, in fact, a 'Cartesian' idea of what a mind is), I go instead a step further, claiming that collectives do feature a collective mind inasmuch as an informational view of mind is endorsed and, then, collectively shared exograms support the collective memories of the integrant of the group.

In the third section, I drew on a certain interpretation of Vygotsky's social dimension of the mind: I underpinned my claim of a collective mind as a shared scaffold whose cognitive dynamics are supported by the collective use of representational artefacts. Collective minds emerge from the scaffolds of the cognitive ecologies in which communities live.

Finally, in the fourth section I concluded that the fundamental scaffolds by means of which communities organize and manipulate their shared exograms (their collective representations) are public narratives.

In the next chapter I shall explain which kind of dynamics permit the emergence of the structures of an ecological niche within which, then, these representational artefacts are built, manipulated and shared. I shall also explain which kinds of information are relevant in the niche where the collective mind of a community emerges.

CHAPTER V

Information into the niche: structures, norms and emergent dynamics

I introduced in chapter III the concept of 'ecological niche' as it is treated in the current literature: namely, the physical adapted space a determinate agent builds all around her, in her habitat, which consequently represents a conditioning factor for the future evolution of that very organism. This is the Niche Construction Theory. This concept is crucial in my thesis, because the collective mind approach I am arguing for entails the extension of the agents' cognitive activities into the surrounding environment and such surrounding environment is, after all, an ecological niche. Moreover, I extensively explained that agents acting in a certain niche produce an *ecological inheritance*, which makes possible their indirect relationship and ultimately results to be a main selection factor.

In that chapter I also remarked, drawing on some of the more influent theorists in the field, that niche construction is not merely limited to the narrow range of the physical structures an agent builds in her environment; rather, depending on the skills of the considered species, niche construction may involve both social and even cultural elements. Drawing mainly on Sterelny's works, then, I defended that the symbolic thought is enabled in humans by the previous presence of a material culture (see *supra*, III, § 2.1); such material culture, as Sterelny claims, supports the emergence of cognitive scaffoldings into the human niche which justify the argument in favour of the human mind as a *scaffolded mind*.

These scaffoldings are mainly made up of artefacts. This is why I devoted chapter IV to the question of artefacts, their cognitive aspects and their relation with the thesis of the collective mind I am arguing for: artefacts support the emergence of collective cognitive dynamics into a human group, operating in the cognitive scaffold of such community. Nevertheless, artefacts are not themselves the cognitive scaffold of a certain community; rather they allow its emergence in the form of a collective epistemic structure: the group's *narrative* (see *supra*, IV, § 3 and § 4).

So, I claim that a collective narrative emerges into a community and co-opts artefacts into collective cognitive processes which are eventually attributable to the whole group. These cognitive processes consist in the production of collective memories and beliefs.

But which kind of information is available in the niche? Clearly, ecological information would seem the most appropriate response: an ecological niche affords ecological information to the organisms living in it. However, it is not so evident what that label refers to when we consider the case of the *cultural* niche: do cultural structures afford anything? If so, do they afford something in the same sense as the physical structures of a certain habitat do?

In this chapter I shall deal with these questions. In the first section I shall analyse the difference between affordances and constraints, and I shall also distinguish these elements from conventions. In the second section I shall address the issue of emergence of structures of information into the niche, with a particular focus on the cultural niche. Finally, the third section will be devoted to analyse the issue of ecological information into the cultural niche and its relation with *representations*; an attempt of reconciliation between representations and the ecological approach will be offered; eventually, I shall explain how representations shared into a certain cultural niche can acquire a normative value and they may carry out a self-regulative role in the community, implementing dynamics of mindshaping (Mameli 2001).

1. Affordances, constraints and conventions

'Affordance' is maybe the most representative theoretical element which distinguishes the ecological approach to psychology as an alternative to cognitivist theories about the nature and the processes of cognition, in order to overcome the long-standing dualist perspective which separates the subject and the object of cognition.

The term 'affordance' was coined by James J. Gibson to define some specific features of the environment an agent lives within: therefore, «[t]he *affordances* of the environment are what it *offers* the animal, what it *provides* or *furnishes*, either for good

or ill» (Gibson 1979, p. 127). Then, affordances appear to be peculiar features of the environment which have to be considered relatively to a specific animal active in that very environment. Affordances are not just resources of the environment an agent can take advantage of; rather an environment might even afford a danger for its inhabitants, such as a cliff affording 'falling' to a non-flying species.

So affordances refer to some specific *possibilities of action* available for a certain organism in its niche (Heras-Escribano 2019, p. 3); they do not consist in mere structural properties of the environment. As Gibson clarifies, affordances «have to be measured *relative to the animal*. They are unique for that animal. They are not just abstract physical properties. They have unity relative to the posture and behavior of the animal being considered» (Gibson 1979, p. 127). In this sense, affordances have a relational nature, because they exist in the relation of a determinate agent who deals with her environment.

So, as Gibson defined this concept, affordances unequivocally refer to the relationship between an agent and her environment. But an environment is not necessarily made up of concrete materials; in fact, an environment can also be a social one. Hence, the constituents of a social niche, such as other people and their behaviour, afford possibilities for action to determinate agents:

> «[b]ehavior affords behavior, and the whole subject matter of psychology and of the social sciences can be thought of as an elaboration of this basic fact. Sexual behavior, nurturing behavior, fighting behavior, cooperative behavior, economic behavior, political behaviour – all depend on the perceiving of what another person or other persons afford, or sometimes on the misperceiving of it» (Gibson 1979, p. 135).

Even in this case, social affordances exist only because a certain social element is perceived and used by a concrete agent *in a concrete physical space*. Other people make part of the niche and what they do – their behaviour – makes part of the niche as well (see *supra*, III, § 2).

Affordances seem then clearly relational properties even though they entail a certain degree of 'materiality': namely, they are available in a concrete material space, an ecological space; their 'materiality' derives, then, from being available as a consequence of the agent's use of a determinate *ecological information* available into her niche. Ecological information is a feature of the environment which consists in all the invariant patterns of physical surfaces available within a certain niche and revealed to the agent by the ambient light; affordances, hence, are perceived by a certain agent just when she 'picks-up' this ecological information from her particular point of observation (Heras-Escribano 2019, pp. 32-38). In this sense the affordances consist in the possibilities of use a certain agent has of that very ecological information which the ambient light reveals to her.

Finally, ecological information is *specific information*. This means that it is always nomologically related to the same affordances anytime the same agent engages with her surroundings: the surface of a certain chair, made explicit for the agent by the ambient light, always affords the same information from that point of view to that very agent, with that concrete body and that determinate mechanical possibility of use of such information.

Specificity is what makes ecological information relevant in any scientific sense: a concrete pattern always consists in a concrete ecological information, which always provides a certain agent with concrete and specific information. This is why we always perceive the same affordances when we pick-up the same ecological information (Heras-Escribano 2019, pp. 41-42).

So far affordances seem to be about the relations of a concrete agent with concrete *spatially extended objects* available in her niche. This is the same as saying that an organism living in a certain ecological niche deals with the physical structure which surrounds her, like the beaver whose actions are related to its dam and the pond it lives beside. In this framework, the ecological information appears to me as mainly referring to objects, structures and spaces existing in the niche (because the ambient light reveals the *surfaces* to the observer).

Nevertheless, as Gibson says, also behaviour affords behaviour. In effect, the specific behaviour of the agents operating in the niche is itself an element of the niche, of a *social* niche (see *supra*, III, § 2). This consideration originated a specialists' discussion about the nature of affordances, whether they are normative elements of the niche or not (Heras-Escribano and de Pinedo 2016; Heras-Escribano 2019, chap. 4). Namely, considering that behaviour can be analysed through a normative framework as a set of practices an agent follows, are the affordances of behaviour to be regarded as social (and, then, normative) elements? From this point, a misunderstanding originated: if picking up the affordances of the behaviour consists in following a norm correctly, maybe picking up ecological information in general consists in following a norm as well, so that for instance 'going up the stairs' would be something that an agent can do correctly or incorrectly.

Therefore, some theorists (Chemero 2009; Rietveld and Kiverstein 2014; Bruineberg, Chemero, and Rietveld 2018) defend that affordances are a guidance for our ecological behaviour and, then, it is legitimate to include into the range of affordances also those elements which define human social and cultural niche, which concern 'higher-order' cognition and act as a guidance for behaviour as well: that is, norms. These theorists draw on Gibson's statement that «[w]hat other persons afford, comprises the whole realm of social significance for human beings» (Gibson 1979, p. 128), namely, in their opinion, *all* those 'things' our social world is made up of reveal to an agent a certain range of *social* affordances. In this sense, they claim that what an agent finds in her niche is fundamentally *general ecological information* (Bruineberg, Chemero, and Rietveld 2018), that is information of general kind about the possible behaviours in a certain environment. This includes both physical information about the surfaces and spaces the agent deals with and more general non-lawful information about cultural conventions and social constraints. So, cultural conventions and social constraints manifest some 'regularities' (Bruineberg, Chemero, and Rietveld 2018, p. 5233) which, in the opinion of these authors, guide our behaviour in ecological terms:

> «*General ecological information* is any regularity in the ecological niche between aspects of the environment, *x* and *y*, such that the occurrence of aspect *x* makes the occurrence of aspect *y* likely. Because of the regular relation between the aspects of the (sociomaterial) environment *x* and *y*, general ecological information allows an animal to couple to a distal (i.e. not sensorily present) aspect of the sociomaterial environment. General ecological information pertains to the ways in which aspects of the environment tend to occur together, like smoke and fire, an object and a shadow, or a pub and beer» (*ibidem*, p. 5237).

However, the kind of regularities these authors point out seem to me to be quite similar to Peircean *indexical relations* (e.g. "smoke indicates fire" or "a beer indicates a pub") and this seems to separate quite harshly the lawful relations based on ambient light information from general ecological information perspective: firstly, these are not necessary (so, non-lawful) relations, but merely *probable relations*; secondly, indexical relations are always *mediated* cognitive relations, so they imply the use of representations (e.g. *abductive inferences* connecting the two extremities of the relation), while the first aim of ecological psychology is to overcome any mediated, cognitivist (then, representationalist) approach; thirdly, they take into account also *distal* aspects of the environment, while Gibson's definition of ecological information seems rather to refer to sensorily-present features of the environment. In fact, Gibson stresses that «[t]he basic properties of the environment that make an affordance are specified in the structure of ambient light, and hence the affordance itself is specified in ambient light» (Gibson 1979, p. 143). So, an observer can directly get all the information she needs about the environment from the environment itself, and such information is trustworthy just because of the nomological relation it is based on.

It is, then, easy to understand why some other thinkers are far more prudent in admitting sociocultural affordances, because their lack of lawfulness would likely trivialize the very notion of affordance (Heras-Escribano and de Pinedo-García 2018, p. 8), compromising the epistemic reach of the ecological approach. In my opinion, this is not an unfounded worry, if we consider that the sociocultural account conceives affordances in terms of *constraints* (Bruineberg, Chemero, and Rietveld 2018, p. 5233)

such as in the case of a beer logo on an aluminium can allegedly constraining its possible content (*ibidem*, p. 5242) and then, in that account, it is a natural conclusion to conceive the beer inside the can as soliciting the observer, albeit it is not directly perceived (being it a *distal* stimulus), through the beer logo (*ibidem*, p. 5241). In particular, this last example does problematically not fit in the ecological account of the direct perception of affordances; instead, it seems to require the use of some kind of representation mediating the perception of the beer inside the can.

In the same way, it sounds kind of weird to say something like "the clock affords me that it is time to take my train" (*ibidem*, p. 5242) or "this landscape affords me to name it 'a cliff'" (Rietveld and Kiverstein 2014, p. 344). Particularly, this last example has been claimed as an instance of the normative nature of affordances: namely, a landscape affords us naming it 'a cliff' because we live in a community which follows such a norm, so that the very norm is the affordance. But considering affordances as normative relations would compromise the direct character of the ecological approach (Heras-Escribano 2019, p. 92). Instead, it seems to me closer to Gibson's original view to name 'affordance' only those practical possibilities of use an agent *actually* has. Each time an agent moves within her niche she perceives many different and even contrasting affordances, but selecting a certain affordance instead of another is certainly a matter of normativity: for instance, a soccer player may ignore the 'scoring' affordance to make an assist to a teammate with more possibilities to get the title of top scorer (Heras-Escribano and de Pinedo 2016, p. 587; Heras-Escribano 2019, p. 109); this is because she is following a determinate norm valid into a particular social background. So norms, intended as social practices, enable agents to prefer a determinate affordance instead of another, but they do not *constitute* affordances as such.

Considering this context, it seems to me more natural to use the term 'affordance' in its proper literal sense only when talking about the ecological niche and the physical 'encounters' an agent may undergo in there. Therefore affordances are the most relevant kind of information an agent deals with into her ecological niche.

But, what about the social niche? I have already remarked, drawing on some of the most recent works in the field, how problematic could be the use of the term

'affordance' in relation to the social niche: someone might in fact think of the affordances as social structures. For instance, theorists who appeal to the concept of 'landscape' to connect the social and the ecological niche trivially misunderstand the relationship existing between social norms and anthropic landscape features: «[e]xercising an ability can be better or worse, adequate or inadequate, *correct* or *incorrect* [my emphasis] in the context of a particular situation, hence there is a normative dimension to the abilities for picking up affordances […]» (Rietveld and Kiverstein 2014, p. 326). The point is that they fail to see that picking-up affordances does not consist in a normative practice, a public rule that you can follow or not and, then, you can do it correctly or incorrectly. It sounds to me a kind of unnatural to say that «my stepping behavior [is] incorrect» (Rietveld and Kiverstein 2014, p. 333): it is maybe *unsuccessful*, but certainly not incorrect. In fact, when an agent engages with affordances, she is not following a rule; she is rather *reacting* to the ecological information she picks up in the ambient light. Therefore, the ecological approach to cognition is, in my opinion, an alternative proposal with respect to cognitivism inasmuch as it focuses on *basic cognition* (Heras-Escribano 2019, chap. 6).

Considering this point, when Gibson affirms that behaviour affords behaviour, I think that he is clearly thinking of behaviour in terms of all the concrete material changes some agents introduce into the ecological space of a different agent through her actions, modifying the surface patterns available in the niche and then the relative affordances; moreover, agents are embodied entities, so their bodies too manifest determinate 'surface patterns' and any action they undertake exists in a concrete physical space and, so, it affects the ecological information of the ambient light.

My point is that it is maybe legitimate, in some sense, to talk about social affordances if we consider social ontology (e.g. laws, costumes, institutions…) as a set of ecological elements with a concrete materiality (e.g. a costume is the concrete practice of that costume; a law is the concrete application of that law…). This would imply that changes in social ontology are after all changes of particular patterns into the ecological niche and, then, they manifest a quite close relation with the rest of the ecological information we perceive in the ambient light of our niche.

Nevertheless, I still feel that something jars with considering sociocultural information in terms of affordances: perceiving an affordance means getting some *direct* information from the environment, but it is not clear in which sense we are perceiving the affordance 'catching the bus' just looking at the clock. In this context, speaking of affordances sounds almost metaphoric because it is not the ecological information available in the ambient light what makes us get out and catch the bus[1]. In fact, the relation [clock-marked time → 'the bus is leaving'] is indeed a non-lawful relation (nor a direct one); it is instead a *conventional* relation (Golonka 2015, p. 243). In this framework it seems to me a bit difficult to keep grounding social ontology and the social niche patterns *only* in the concepts of affordance and ecological information.

This issue becomes even more problematic when considering the case of the cultural niche: in fact, representational elements like *words* or *figures*, which are core-components of the cultural niche a human group inhabits, do not (literally) afford anything. Ambient light reveals properties of the surfaces available in the environment, but words have no surface to reveal; while figures, albeit they do feature a surface, are rather conventional elements of the niche which do not manifest any lawful connection – in their use – with the mechanics of our human bodies. Even though, all the representational elements which constitute the cultural niche are still pieces of information available into an ecological space: representations, in effect, are always represented and displayed on some kind of *support*. Therefore, they indeed make part of the information an agent can take advantage of in her niche. But, if they do not provide affordances to the agents which use them, which kind of information do they provide? I claim that they display *structural properties*, such as constraints, which are intrinsic properties in the same sense suggested by Norman (see *supra*, IV, § 1).

[1] Perhaps I could say that I perceive this or that *linguistic affordance* in the sense that, depending on my actual social condition, there is only a limited range of speech acts (then, possibilities of linguistic actions) available for me at a certain moment in a certain context (Ayala 2016); hence I could consider that such speech acts affect my own behaviour in my niche and the behaviour of my fellows too. Nevertheless, it is not clear at all in which sense the way these speech acts affect my behaviour maintains a *direct relationship* between an ambient factor, such as the phonemes I hear, and my subsequent behaviour: it seems more reasonable to say that what stimulates my reaction is the very *meaning* of the utterance, while the unique ambient factor I perceive is a sequence of phonemes, which indeed *mediate* my understanding of the meaning of the utterance.

As Norman remarked in a short and sharp article more than twenty years ago (1999), affordances, constraints and conventions are not synonyms. Considering the very case of representations in the framework of computer design where they are displayed on a screen, he highlights the most evident problem which extending the ecological terminology into the cultural niche entails:

> «The cursor shape is visual information: it is a learned convention. When you learn not to click unless you have the proper cursor form, you are following a cultural constraint. Far too often I hear graphic designers claim that they have added an affordance to the screen design when they have done nothing of the sort. Usually they mean that some graphical depiction suggests to the user that a certain action is possible. This is not affordance, either real or perceived. Honest, it isn't. It is a symbolic communication, one that works only if it follows a convention understood by the user» (Norman 1999, p. 40).

So we have here the central point: the shape of a representation on a screen is indeed an instance of *visual* information; however, this does not imply that it affords us any *ecological* information. On one hand, it is true that a certain representation *suggests* to the user some concrete actions and *limits* some others; on the other hand, this 'suggestive character' of a representation is regulated by a cultural convention.

It is then a matter of formal *constraints* (not affordances!) which induce a user to manipulate a certain representation in some ways rather than in others. In fact, any representation features some intrinsic properties (see *supra*, IV, § 1) which make it more appropriate to represent a determinate state of the world instead of another. Not everything goes, of course. Representations are not arbitrarily constructed, conventions are (Norman 1999, p. 42). As I shall detail in the next chapter, in fact, the symbolic connection between a certain representation and its meaning is surely conventional; nonetheless, that representation always has to respect some determinate formal constraints which limit its *appropriateness* to its referent, in Norman's word, or better its *iconicity*, in semiotic terms[2]. There are some evident formal constraints which allow a

[2] Remember here the case of the representational appropriateness of the map I spoke of in the last chapter which Norman remarked in his seminal work (Norman 1991, pp. 30-31).

stick to represent a horse, while impede it to represent a ball or a boat (see *supra*, IV, § 3).

In sum, what I claim is that the kind of cognitively relevant elements an agent engages with into her niche are affordances insofar as she deals with ecological information, while they are constraints and conventions insofar as she deals with representations into her cultural niche.

In the next section I shall offer a model to explain how some structural constraints may develop into the ecological niche and how the very representations emerge in there. I shall also explain in which sense these emergent structures, once they have emerged, can feed back to the agents operating in the niche, orienting their behaviour as normative factors.

2. Stigmergy: the dynamics of emergence

I introduced the notion of stigmergy in chapter II: I described it provisionally as the basic dynamics which allows the emergence of hetero-directed coordination into groups of individuals, *mediated by traces*. I claimed that stigmergy is, after all, the general dynamics of self-organization and it eventually is the basic dynamics of the feedback loop which develops between an agent and the niche she lives within. Stigmergy, then, is a mechanism which can help to analyse the emergence of concrete structures into a certain niche, whether ecological or cultural.

As I mentioned in chapter II, 'stigmergy' is a neologism coined by the French entomologist Pierre-Paul Grassé at the end of the 1950s (Grassé 1959): meaning 'work from a stimulus' (*stigma* + *érgon*), in the mind of this scholar it would solve the enigma of coordination in groups of social insects, whose individual components are supposed to be too simple entities to carry out complex cognitive operations alone, such as building the intricate architectures of their nests. The solution proposed by Grassé was that no one of those simple agents (termites, in his prototypical case) was aware of the 'project' of the nest; their high pillars were a stochastic emergence from a handful of simple algorithms individually followed by each agent. In fact, each termite accumulates pieces of mud

where she founds other mud marked with pheromones as an incipient deposit of building material – the 'grounds' of the pillar – like I illustrated before in Fig. 2.1 and in Fig. 2.2. So each termite's actions were stimulated by some kind of *structure* or *trace* left in the common work-space by another agent. The conclusion is that the very 'memory' of such system (the tasks which have already been done and those which have not been done yet) is *distributed* into a concrete work-space, while agents' interaction is *indirect*, mediated by their environment (Bonabeau et al. 1997, 1998).

This principle of stigmergy took long time to expand to other fields of research (Theraulaz and Bonabeau 1999). Only at the beginnings of the 1990s the concept of swarm intelligence was introduced into computer science in form of agent-based models to study group cognition (Heylighen 2016, p. 5). Then, virtual simulations started to be reproduced in robotics: ant-like robots were programmed with simple algorithms to carry out a determinate task getting cues from the environment, such as recollecting and gathering objects in some concrete points of the work-space (O. Holland and Melhuish 1999). It is only in the 2000s when the concept of stigmergy started to be applied to human instances of group cognition (Doyle and Marsh 2013): emergent patterns in economics, for instance, started to be explained as the self-organization of groups through stigmergy, where the environment was the market, the agents were the sellers/buyers and the traces were price oscillations (Heylighen 2016, p. 7). Hence, some authors noticed that even Adam Smith's well-known metaphor of the 'invisible hand' organizing goods allocation into the market is, after all, a case of stigmergy in a human context (Marsh and Onof 2008, p. 140).

2.1. General features of stigmergy

The phenomenon of stigmergy features the following five core-elements: a) *action*, namely any change introduced in the state of the world with respect to its previous status; b) *agent*, an autonomous system (Kauffman 2000) oriented towards an objective described by the algorithms responsible for its behaviour; c) *condition*, the state of the world which stimulates a reaction from the agent; d) *medium*, the environment shared

by the agents which constitute the group system, their work-space; e) *trace*, the change introduced into the medium by an agent's action. The medium is necessarily shared by all the agents taking part into the system; it has to be accessible and manipulable for all the agents involved in the dynamics of stigmergy; the medium *mediates* the interaction between the diverse agents operating in there. The medium features a certain set of conditions which stimulate the agents' behaviour; any condition may stimulate but never determine the reaction of an agent, because it is just a sufficient but not necessary condition of her reaction, which eventually is merely something likely, not expected. Each agent always does something to get her objective and every action she carries out necessarily changes the previous state of the medium, leaving a trace of her action into the shared environment. Any trace is therefore a consequence of the agent's action and then it carries some information about such an action: it is a 'message' which imposes a cognitive challenge for the other agents (e.g. a new problem-solving strategy). Finally, any trace entails a new condition into the medium, which stimulates new reactions from the other agents. So stigmergy describes the feedback loop existing between different agents sharing the same work environment (e.g. the same habitat) and it can be defined as follows: «stigmergy is an indirect, mediated mechanism of coordination between actions, in which the trace of an action left on a medium stimulates the performance of a subsequent action» (Heylighen 2016, p. 6).

Hence, the dynamics of stigmergy is a model to explain the emergence of coordination in a group of many agents who mainly do not maintain direct relations among them. This is why it is a good candidate to explain the mechanism of niche construction and the emergence of new structures in the niche as well[3] (see *supra*, III, § 1.1): in fact, the habitat shared by different organisms could be considered as a medium where all the agents (all the organisms) interact with some conditions they find in the medium; they can perturbate the medium consuming resources, producing waste or building structures such as their burrows and all these operations *leave a trace* into the medium. Such trace is not necessarily something like a footprint left on the sand; in

[3] Disclaimer: obviously, stigmergy just explains *the emergence of new structures* into the niche through the feedback loop agents/environment; it does not directly concern the evolutive relevance of the environmental changes for the species which inhabit a certain niche.

effect, also a structural modification of the habitat, such as the construction of a beaver dam, is to be considered as a trace of a previous concrete action operated by a concrete agent (in this case, a beaver) into the medium. Each trace results to be a new condition in the habitat which all the agents have to deal with (certainly, each one in her own way), eventually triggering new actions.

The importance of stigmergy therefore is mainly due to its character of *asynchronous interaction*: in fact, in this framework there is no need for the agent's actual presence when her trace produces its effect on another agent: traces are relatively persistent in the medium (Ricci et al. 2007, p. 132; Heylighen 2016b, p. 53), so coordination can be diachronic and the very emergence of new structures is diachronic.

So far I have introduced the concept and the standard features of stigmergy as a general principle for emergent coordination and self-organization into a group of agents. Nevertheless, in this general presentation I have considered a basic form of stigmergy whose agents are conceived as simple units, ant-like entities as in a standard multi-agent system (Ferber 1995). Indeed, this is not the unique context of application for stigmergy; its epistemic reach is far wider. In fact, even many collective activities of human groups can be treated in terms of stigmergy and I shall focus on this issue in the next section, where the stigmergic emergence of new structures in the human cultural niche will be tackled as well.

2.2. Human stigmergy

More recently, an increasing number of theorists noticed that the agents operating in a multi-agent system do not have necessarily to be conceived as very simple units like in an ant-like system, where the environment is oversimplified as well (Marsh and Onof 2008; Lewis and Marsh 2016). They extrapolate the basic ideas underpinning general stigmergy, to apply them in very different and quite more complex cases involving human agents interacting in social environments. In this sense, some scholars proposed new labels such as *cognitive stigmergy* (Ricci et al. 2007; Omicini, Ricci, and Viroli 2008;

Lewis 2013) or *human stigmergy* (Parunak 2006; Doyle and Marsh 2013; Lewis and Marsh 2016) to describe the new possible contexts of application of this theoretical paradigm.

The main point in this new cognitive focus is that agents do not undertake tasks just following simple algorithms; they are instead intentional agents with concrete objectives (Ricci et al. 2007, p. 129) and environments are not elementary work-spaces featuring mere pheromone-like traces (Tummolini and Castelfranchi 2007). So, in this new framework the environment is conceived as an articulated work-space where psychologically complex agents manipulate a lot of different artefacts; it is not understood as a mere passive landscape anymore (Ricci et al. 2007, p. 128). Examples are now taken from real social life to highlight how people change the physical structure of their immediate surroundings any time they carry out some actions: for instance, Mary and Julian are having lunch; if Mary gets the salt, she is redesigning the table-environment in such a way that Julian himself cannot get the salt; Mary is implicitly saying "I am using the salt, you cannot use it right now". Notice that this amounts to saying that Mary is redesigning their environment (Kirsh 1996) in a way that limits the affordances available for Julian at a certain moment; this is because she is actually 'modifying the niche' in her immediate surroundings. Hence, in this example, the medium/environment is the table where Mary and Julian are sitting, while the traces are the changes in the relative positions of the objects on it.

Certainly, human social world features plenty of similar and more complex cases, especially when people share the same space and resources in situations of co-working. Both real or virtual situations invite people to get involved in instances of stigmergy, where coordination is possible through the manipulation of the medium: in remote software and web-page co-editing, for instance, the medium is virtual and coordination emerges both from the very modifications introduced in the code by each programmer and from specific *annotations* (the human equivalent of pheromones) left here or there by this or that co-worker (Bolici, Howison, and Crowston 2016); in a real space, instead, coordination emerges from modifications in the concrete co-working environment, such as changing the position of some artefacts or manipulating them (Christensen 2013). The two most representative examples, in this last case, are the control-panel based model

and the blackboard/Post-It model (Ricci et al. 2007; Heylighen 2016b; Susi 2016): in the first case, all the workers of a certain factory have access to a common room where a control panel represents the current state of the tasks to be carried out in the factory by means of its different layers (e.g. on/off switches, lights, etc.), so that any worker can at any time access a real-time representation of the current state of the work, what has already been done and what task still needs intervention and she can act consequently (Susi 2016, pp. 43-44); the second case is even more common and we can, for instance, think of an office where an employee finishes her task and passes over the work to other co-workers dropping off, let's say, a pile of forms on a table with a Post It on them saying "Please, fill in!" or "Please, photocopy!" (Heylighen 2016b, p. 55). All these modifications of the work environment (the shared medium) are eventually *traces* which mediate the interaction among the integrants of a determinate human group.

Finally, stigmergy also works for the very individual agent who leaves a trace in a medium: that very trace is not just a condition for new actions operated by other possible agents sharing the same medium; instead, it can stimulate a reaction in its very author as well. In this case we can speak of *individual stigmergy*. This is indeed a very common phenomenon in individual creativity. A concrete example could be a sketch or a clay model which an artist outlines and then modifies and moulds: «This preliminary registry of the work performed calls out for more. It challenges the user to add, to enhance or to correct. Each addition changes the trace, thus attracting the attention to further imperfections, or suggesting further additions» (Heylighen 2016b, p. 55). Having a model clay shape is crucial for, let's say, a sculptor who is working on a new idea: manipulating such artefact, an artist can refine some details or get new insights from it; it is after all a particular *epistemic action* (Kirsh and Maglio 1994).

Stigmergic prototyping (Kiemen and Ballon 2012) – as this is called – is not necessarily an individual activity. When sketches are shared in communities of artists, for instance, creativity becomes an instance of multi-agent stigmergy whose unpredictable emergent outcome is the final piece (Secretan 2013, pp. 69-70). In this sense, « 'public traces' » (Heylighen 2016b, p. 57) stimulate the responses of the agents who share the same medium, for instance, the same lab or the same art gallery.

This last point is crucial for my central argument because public representations are just this kind of 'public traces' left into a shared environment, left into a shared cultural niche. At the end of this section it results clear that the emergence of a public representation, with all the constraints and conventions related to it, can be treated just as a peculiar case of collective stigmergic prototyping. Let's think, for instance, of a certain icon: the same subject may inspire many different representations of it, each one featuring its peculiar formal properties (e.g. its particular dimensions, its colours, its particular materials, etc.); when all these representations are shared in the same cultural niche, they challenge all the integrants of the community to enhance or to correct them; ultimately, a prototypical representation of that subject they all refer to emerges. I claim that such a collectively-built prototypical representation becomes a model with a normative status for the community that produced it: a *canon* which regulates the future relations of reference with respect to that specific subject.

In the next section I shall focus on the structures of a cultural niche; on one hand I shall distinguish between the *ecological information* concerning culture and, on the other hand, *representations* as typical elements of the cultural niche. I shall finally offer my interpretation of the role of representations into the economy of human cognition.

3. Cultural structures and their meaning

So far I analysed the different kinds of information human agents find in their niche: I remarked the difference between affordances and ecological information, on one hand, and constraints and conventions on the other hand. Then, in the last section, I analysed the mechanism of stigmergy as a model to explain the emergence of new structural features into the niche, both ecological and cultural. I also highlighted that emergent structural patterns can *reorganize* the behaviour of the agents belonging to the same system. My objective, now, is to clarify in which sense representations are emergent structures into a cultural niche and which role they carry out with respect to human cognition. But, first, I have to consider some recent attempts to treat cultural products in ecological terms.

A valuable contribution to the study of human cultural niche recently came from a series of new studies, mainly authored by the anthropologist Tim Ingold and the archaeologist Lambros Malafouris. They proposed an innovative focus on material culture for research about human cultural life. Concepts like *ecology of materials* (Ingold 2012, 2013) or *material engagement* (Malafouris 2013, 2019) redesigned the contours of the object of research, remarking that we humans are *embodied* entities living in our own peculiar niche. So, if cultural elements have a meaning, they have it as elements of our ecology and inasmuch as they are related to our embodied life.

So, against the Aristotelian hylemorphic biases characterizing the received view both in anthropology and cultural studies, Ingold (2012, p. 432) stresses the importance of the very materials we use in the construction of our material culture: for instance, the clay we use to make bricks is not a merely passive recipient of an ideal form an agent communicates to the matter, as an hylemorphic view would hold; rather, that very clay manifests some peculiar properties (e.g. being more or less malleable or resistant) which affect the way in which the agent gives it a shape (Ingold 2012, p. 433; Eco 1997/2016, p. 72). Therefore, also materials play a determinant role into the forging of human material culture. This is possible because human agents certainly are embodied entities, living within an ecological space which ultimately affects their cognitive behaviour (Ingold 2007).

Lambros Malafouris maintains as well that human mind is strictly interwoven with material culture and the use that we make of artefacts in our cultural life. He stresses the embodiment of human cognition and, at the same time, he remarks the centrality of an ecological account of our cognitive skills. Similarly to Ingold, his point is that when instances of material culture are produced, the final object is not a mere realization of an alleged mental representation which pre-existed the creative operation: for instance, in the prototypical case of a potter producing ceramic vases, the artisan is not merely applying a mental model to an inert matter such as the clay may be conceived through a cognitivist bias; instead, the potter's hand movements are also guided by the very emergent affordances of the clay, so that the final form results from the convergence of determinate hand movements and the resistance opposed by the very

material volume that the artisan manipulates (Malafouris 2019, pp. 9-12). The centrality of the hand and manual production in human cognition is highlighted by Ingold as well (Ingold 2013, chap. 8), so that human cultural production is a clear outcome of an embodied mind, situated into a certain ecological niche.

I do subscribe this perspective of a manifest continuity between the ecological niche we inhabit through our embodiment and, on the other side, that peculiar trait of human ecology which is the cultural niche. In fact, as I have been arguing along chapters III and IV, symbolic thought and representational culture emerge in my opinion on a pre-existent material culture. In this sense, I see perfectly coherent to analyse, in an ecological framework, the relationship which connects the typical artefacts of a concrete material culture with determinate production techniques and the consequent uses a human community attributes to them.

However, a difficulty arises here: which role does correspond to representations in this context? In fact, it is perfectly acceptable, for me, to explain the concrete use a certain artefact might have been thought for in a determinate community just interpreting the creation process as a dynamic relationship, where the material to be transformed, the affordances it features and the embodied skills of the artisan converge. This is evident in the example of the pottery production mentioned by Malafouris (Malafouris 2019, p. 11). Nevertheless, the framework proposed by both Ingold and Malafouris just considers cultural artefacts insofar as they are conceived as material elements, as concrete pieces of a certain material culture. But a cultural niche does not merely consist of those elements its correspondent material culture is made up of. Material artefacts and tools such as pottery or hammers and axes can certainly be studied in terms of their ecological role for human embodied minds, because they feature a particular ecological information which affords all their possible uses compatible with a human body. But, as I argued before (see *supra*, V, § 1), a main part of our cultural niche is made up of representations (e.g. words and figures) and such representations do not feature any affordance at all. They just feature constraints and conventions. Therefore, they cannot be treated in the same ecological terms as we do with a ceramic vase or a stone axe. So, how should we approach them? Firstly, it is necessary to clarify

what I mean for representation and this is a public shared instance of a representational element such as a word or a figure; then I would consider as structural emergent elements of the cultural niche the constraints which a concrete representation features. Hence, I suggest that the best way to approach them is a semiotic analysis of the artefact which supports them in order to detect the relative constraints.

As I said before (see *supra*, V, § 1) representations are still elements we experience in our physical surroundings, as a part of our ecological niche; this is because they are always displayed by means of some concrete artefact (see *supra*, IV, § 1 and § 3). Nonetheless, we cannot understand their value within a certain cultural niche merely appealing to the ecological information afforded by their physical support. Instead, representations manifest, firstly, some intrinsic properties in form of graphical *constraints* which make them more or less 'appropriate' (see *supra*, IV, § 1; V, § 1) to represent their referent. As I shall argue with more details in the next chapter, representations always respect some *formal requirement* (iconicity, in case of figures) which makes them effective when representing. Secondly, representations show some symbolic features which, despite having some close connection with the formal properties of the representation, are indeed *conventional* elements.

Formal properties – the iconic features of a figure – can emerge and evolve into a certain community through a process of stigmergic prototyping (see supra, V, § 2.2). Even though, they always maintain some clear kinship because they always respect *the same constraints in order to refer appropriately* to their referent. Let's see some examples:

Fig. 5.1 Roman Venus. Unknown author. Pompeii (I century)

Fig. 5.2 ***Sleeping Venus*****. Giorgione (1510)**

Fig. 5.3 ***Venus of Urbino*****. Tiziano Vecellio (1538)**

Fig. 5.4 ***La maja desnuda*****. Francisco Goya (1800)**

Fig. 5.5 ***Pauline Bonaparte as Victorious Venus*****. Antonio Canova (1808)**

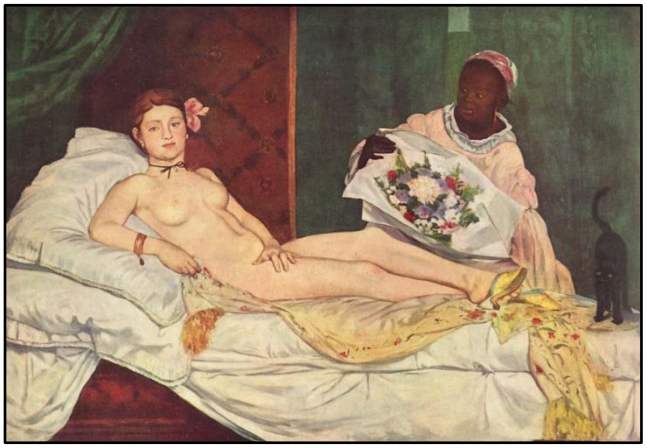

Fig. 5.6 ***Olympia*****. Édouard Manet (1863)**

These figures illustrate an interesting phenomenon: they all feature the same *model*. When we want to effectively represent (that is, to refer to) let's say a Venus, there are

some concrete formal elements that our representation has to respect, just as we have to write a word respecting its orthography if we want other people to understand the signs we have written and, then, the possible meaning they refer to. Notice that here I am not speaking of any symbolic element yet. For now I am just considering the constraints of a certain representation inasmuch as they are formal intrinsic properties.

In the next chapter I shall clarify what iconicity implies and to which extent these formal elements can be considered as intrinsic properties of the representation. For now it is enough to say that an icon is not a 'double', a copy of its referent; it is rather a translation of that referent into a certain *code*, just as a sequence of letters forming a word does not constitute a 'double' or a copy of its referent. In the same sense, I remark that a figure does not *reproduce* a certain object; it just *refers* to that object.

I introduced this point before in my argument (see *supra*, II, § 3.1.2): a graphical representation is a translation into a concrete graphic code of a certain utterance and, specifically, an *ostensive utterance*[4]. When I endorse a certain representation of a Venus, I am literally saying "This is a Venus!". A certain graphic code corresponds to such utterance. If that very graphic code is eventually the outcome of a collective stigmergic prototyping, the correspondent utterance has to be attributed to the whole collective responsible for the emergence of such a graphic code. So, this last utterance whose subject is a collective subject is, ultimately, equivalent to a *collective belief*. This is, in sum, what I have been arguing in chapter II, § 3.1.2.

Nonetheless, in this framework one more point requires to be clarified: what is the role of representations in the economy of human cognition? Representations are generally conceived in the literature about cognition as internal states which mediate the relation between the mind and the world, such as the Fodorian mental symbols which compose the 'language of thought' (Fodor 1975). My account of representations is certainly quite different. First of all, I *do not* regard representations as an intermediate element between perception of sensorial information and brain implementation, mysteriously existing somewhere in the head. Instead, because of my naturalist and

[4] I refer here to the definition of ostension offered by Umberto Eco in his *Trattato di semiotica generale* (Eco 1975/2016, chap. 3.6.3).

ecological focus, I believe that representations have to be explained coherently with this account as constructions *ex post* with respect to perception and cognition. Therefore, I consider that representations primarily are core-constituents of a human cultural niche; they are supported and displayed by representational artefacts such as paintings and the like; they eventually are concrete instances of a certain graphic code. When such a graphic code – which is just a rule of transposition between the representation and its referent, as I said before (see *supra*, II, § 3.1.2) – corresponds to the outcome of a collective stigmergic prototyping, this graphic code is recognised and endorsed as a good act of reference by all the community. When it is endorsed in this way by all the community, it becomes a *canon* and acquires a normative power (inasmuch as it is a canonical representation, the 'right' version) on the agents who constitute such a community, just as an emergent structure in a multi-agent system affects and redesigns the behaviour of the integrants of that system (see *supra*, V, § 2.2).

I claim that such canons, considered as representations publicly recognised and endorsed within a determinate community, become normative stimuli which carry out a crucial meta-cognitive function for the integrants of the community: in fact, they regulate the acts of reference acceptable for the members of a certain group. I would remark that I do not intend, here, 'meta-cognition' to refer to any kind of second-order mental representation or the like. Instead, I mean for 'meta-cognitive function' a function which does not make part, strictly speaking, of the cognitive process, but which nonetheless *conditions* the cognitive process 'from the outside', for instance, it is meta-cognitive the regulative function of folk-psychology (McGeer 2007). I shall contend, in fact, that canonical representations likewise exert a regulative function for cognition into their cultural niche.

Even though, as Norman teaches us, there are two ways to look at a system (see *supra*, IV, § 1): the system view and the agent's view. Hence, in the agent's view, a public representation communicates a canon which regulates the individual agent's acts of reference; we have here the *meta-cognitive function* of public representations. Instead, in the system view, a public representation constitutes a collective belief because it constitutes an act of reference attributable to the whole group; so a public

representation carries out a *cognitive function* with respect to the group understood as the subject of a collective mind.

The regulative (meta-cognitive) function of representations is eventually a form of *mindshaping* (Mameli 2001), that is a moulding activity operated on the mind of a certain subject by something or someone. In the next chapter I shall introduce the concept of mindshaping and I shall argue in favour of a *categorial mindshaping* as the main meta-cognitive function of public representations.

CHAPTER VI

Categorial mindshaping

The aim of this chapter is to analyse and criticize a considerable omission in the contemporary literature about mindshaping: the role of public images in moulding the minds of the others. What I argue is that, whenever we design an *imaginal space*, we are creating *physical* representations of a conceptual content: we are designing *an ecological niche with conceptual features*. With the words 'imaginal space' I define a virtual space constituted by a coherent set of images existing in a determinate cultural niche; in some cases it can be even an actual space defined by a coherent series of real images which tell us a specific story: let's think, for instance, about the figurative representation of a temple telling us the coherent story of a myth[1].

Clearly, we do not design a social niche (Sterelny 2007; Kendal, Tehrani, and Odling-Smee 2011; Laland and O'Brien 2011; O'Brien and Laland 2012; Ryan, Powers, and Watson 2016) by means of verbal practices only; we also do it directly representing a concept through an iconic artefact. In this sense, any artwork can be a vehicle for a concept it iconically refers to. Obvious examples are indeed political propaganda artworks (such as sculptures and paintings representing a ruler) or religious propaganda artworks (like mosaics or frescoes representing God's creation of the universe or the lives of saints). What all these artefacts have in common is that they all represent *how the world is* (descriptive function) and *how it has to be conceived* (normative function) to every member of the community living within such a social niche. I shall claim that it is this double nature of such an imaginal space what produces the mindshaping effect within the community.

The term *mindshaping* first appeared in Mameli (2001), an article where the author presents mindshaping as the niche-construction effect of our mindreading practices. Niche-construction is the global effect of all the actions an agent performs to

[1] See the temple metopes of the Heraion of Foce del Sele, at the Museo Nazionale Archeologico di Paestum (Italy).

improve its fitness in its ecological environment. Therefore, mindshaping is a 'side effect' of our mindreading activity of other minds: considering that our social niche is also composed of mental features and the mind of the others is one of them, to read it more easily and successfully we undertake a lot of 'side actions' which literally *shape* how the others think.

We do it in a lot of different manners: for instance we assume diverse *emotional approaches* in different contexts to help the others to understand us and to be sympathetic with us, or to 'suggest' them what attitude to assume in a particular circumstance; so we show an *angry* attitude when someone hits us, we show a *disgusted* face smelling something stinky, we show a *sad* face when any accident frustrates our plans or we show a *pleased* or *unpleased* face looking at the practices of the others. This is one of the ways we 'teach' the others how they should react in specific contexts.

In the first section of this chapter, hence, I shall shortly resume the mindshaping theory Matteo Mameli presented in his seminal work. Then, I shall analyse in the second section Tadeusz W. Zawidzki's recent contributions to the cause of mindshaping (Zawidzki 2008, 2013): he argues that mindshaping is a primordial cognitive strategy which humans developed during their evolution, a strategy whose appearance is prior to mindreading practises. He focuses both on pedagogy and on a more general tendency people have to transmit practical knowledge to their offspring. Zawidzki pays a lot of attention to the ways people undertake to gather and transmit information: for instance he focuses on the relationship between a master and his apprentice in any technique learning process. He devotes the main part of his recent book to analyse the psychological cognitive features of mindshaping, just mentioning the role a foundational myth plays in giving identity and cohesion to a particular community and in transmitting practical knowledge to the future generations. Then, he simply ignores how we shape the way in which our community conceives itself, managing the physical space it lives within: for example, he does not say anything about how we use architecture to shape the intentions of the others towards a specific space and, definitely, he does not speak about how we use images to manipulate and create conceptual features in an

imaginal space. This is the task I shall assume in section § 4, after having analysed some lacunae of the classical mindshaping approach in the third section of this chapter.

1. Mameli's mindshaping theory

In his paper, Mameli starts considering how mindreading works and how it influences our niche construction practices: we are enculturated and aware of our mental states in a very different way from other organisms; this makes us good mindreaders, able to predict the mental states of our conspecifics and their behaviour (Mameli 2001, p. 597). Mindreading is crucial for the main cognitive capacities characterizing the human species, such as language, epistemic progress, cultural evolution and the very niche construction. So, if signalling is a quite widespread activity in the biological domain, linguistic communication characterizes only humans. The main difference between these two kinds of communication is that in the last one the information receptor attributes the *intention to* communicate to the speaker, while is not the case in the former one. In this sense the ability to attribute communicating intentions to the others is just a peculiar mindreading skill (Mameli 2001, p. 598).

Likewise, epistemic progress entails selecting the best problem solving strategy and this implies the capacity to evaluate both one's own and someone else's beliefs, considering their truth value and their relevance in a specific context, which is another particular mindreading skill.

Also cultural evolution in humans is characterized by a quick diffusion of ideas, such as the use of a specific tool or jewels and clothes: while other animals just imitate (through social learning) some conspecific efficiently using a new tool, humans improve their artefacts applying a teleological framework ("How could I use this object?") to all the possible uses of a given tool. This practice too is a mindreading skill, being it based on the analysis of the intentions and the possible aims a peculiar object can activate in conspecifics.

Finally, we do live in an ecological niche, a space we gradually build all around us each time we perform an action useful to improve our fitness within our environment.

Our mental life is an important part of this ecological niche, and it is composed of all those beliefs we attribute to the others and those of ours we are aware of. Being able to predict (that is, to 'mindread') the others has indeed *mindshaping* effects on us (Mameli 2001, p. 599), changing the mental features of the niche we live in. In this sense Mameli rightly observes that our niche construction practices change the selection pressures operating within our niche, both on us and our conspecifics and descendants; this makes niche construction evolutionarily significant.

Making our minds part of the niche we inhabit implies that, each time we act on the niche, we can act on the others' minds. We do it through *expectancy confirmation*, which results when some object meets the expectations of an observer (Mameli 2001, p. 600). Expectancy confirmation effects are, therefore, *mindshaping effects*. A mindshaping effect is an effect on the very structure of the mind and its development: for instance, when I tell you I've just broken your smartphone, this *makes you angry*; when a father teaches his son how to tie his shoes, this *makes his son learn* how to tie them (Mameli 2001, p. 608). These are both examples of mindshaping effects.

So, considering that niche construction consists in all those actions an organism undertakes to improve its fitness in its environment; considering that, in an environment where minds are a main feature, an efficient mindreading is a necessary trait to improve the fitness of an agent in such an environment; mindshaping seems to be a *strategy* to make mindreading easier. It is in this framework that the attribution of traits and states to someone or about something produces those peculiar mindshaping effects called expectancy confirmation effects.

In his paper, Mameli presents a bunch of quite incisive examples of expectancy confirmation effects. I shall report here just three of them to better explain what mindshaping is:

> a) «A doctor tells one of his patients that she is depressed. As a result she starts looking at her feelings and at her situation in a new light. She starts suspecting that she may really be depressed after all. She sinks into a depression». (2001, p. 609);

b) «A father expects his children to share his own values. The father's expectations put a lot of psychological pressure on the children. As a result of this, the children end up valuing, at least in part, the same things as their father». (2001, p. 609);

c) «A group of people, the W's, think that people belonging to another group, the B's, are violent. Because of this, the B's get to be treated in a special way. Because of this treatment, the B's find themselves in situations that lead them to behave more violently than the W's do» (2001, p. 609).

It is clear, then, that specific attributions of mental states by a subject A (his *expectancies*) to a subject B make A behave in such a peculiar way with B that eventually *shapes* the very mental states of B.

In this context it is necessary to note that expectancies can be transmitted both vertically (from one generation to another) and horizontally (within the same generation), by means of *social learning*. Synthetically, Mameli just focuses on three main social learning ways to 'mindshape': through *teaching*, that is by *explicit transmission* of some knowledge (theoretical or practical) from a teacher to an apprentice; through *imitation*, that is the conscious repetition of a particular action or behaviour until its correct realization; through *non-imitative learning*, that is for instance learning a behaviour by non-intentional transmission, by the observation of a model just like a disposition a child gets by observing his parents' attitudes towards someone or something.

2. Zawidzki's mindshaping approach

While Mameli considers mindshaping as a peculiar niche construction strategy whose main aim is to improve our mindreading abilities as well as our fitness in the niche, Zawidzki points out that it is quite unlikely that mindreading has been selected by

evolution for its (alleged) efficiency in predicting the mental states of our fellows. Instead of mindreading, propositional attitude ascription serves, for him, a mindshaping function: it enables us to set up *regulative ideals* to mould the behaviour of the others (Zawidzki 2008, p. 194). He argues that folk psychology, the attribution to others of a psychology based in a belief/desire propositional attitude, entails a mindshaping aim: namely, an easier coordination with our conspecifics mediated by *rules* our folk psychology is accorded with. Sometimes we elaborate a 'theory' to interpret and to predict the behaviour of the others; sometimes we just apply *simple heuristics* to their actions; sometimes we complement both mindshaping and mindreading strategies to make sense of the actions of the others. However, the most interesting implication is that the mindshaping approach does not consider the coordination task we undertake with our fellows analogous at all to managing our relation with the other elements of our ecological space. This is, we do not just *observe* in a third-person perspective the mental states of the others and then elaborate a 'theory' to explain them; we can *act on* them and change them, because we use a second-person perspective in relation to our fellows (Zawidzki 2008, pp. 195, 201, 204).

The reason why folk psychology (despite what mindreading theorists pretend) is a quite inefficient predictive tool is, clearly, that it is not quick nor reliable enough: the *holism problem* (Zawidzki 2008, p. 195), the thesis that any belief/desire couple admits an indefinite number of possible coherent behaviours, invalidates its supposed predictive power. Hence, considering mindreading as the main aim of folk psychology would be pretty problematic, no matter if in a theory-theory approach or in a simulation theory perspective: in both cases uncertainty would be unavoidable (Zawidzki 2008, pp. 196-198). This is why folk psychology's real importance consists in its mindshaping function. Let's consider, for instance, the city traffic case: in the risky and challenging environment of the daily traffic in a busy city, there is no way I could divine the actual causally relevant mental states which make any of my fellows take a certain direction or a different one just applying a folk psychology paradigm: there, in the crossroad where the two main avenues of the city intersect, I have no time nor means to figure out what is going on in the mind of the driver I have in front of me. So, how could I manage to

avoid the likely crash? It would be a very thorny epistemic puzzle to solve if I could not rely on a well settled *normative system*, like the one formed by traffic signals and driving rules (Zawidzki 2008, p. 199), a normative system that indeed my fellows do not ignore because they have been well socialized to respect it.

So far, it results pretty clear why mindshaping is an effective and efficient coordination strategy. The reason resides in the *normative force* of the expectations we have of our conspecifics: we know the normative system, they know it too; they do *have to* behave following the rule R in the context C, while the rule R' in the context C'. Mindshaping's strong point is, therefore, that it is focused on *prescription*, not on *prediction* (Zawidzki 2008, p. 202). Being aware of the normative system prevents us from acting erroneously avoiding, thus, the correspondent sanction. This means that we can interpret the behaviour of others considering what they *have to* or *do not have to* do in terms of an external system of rules. A system of rules that also for Zawidzki, like already suggested by Mameli, is transmitted by explicit or implicit teaching or learned by imitation (Zawidzki 2008, pp. 205-206).

Furthermore, in a recent work (2013), Zawidzki focuses more organically on mindshaping as our socio-cognitive linchpin. Again, he argues that, in an evolutionary framework, it does not make any sense to focus on mindreading typical propositional attitude attribution as a cooperation factor. He suggests that, in any case, this would be at most a later evolutional step, while a prior one would be indeed the implementation of mindshaping strategies. Within this framework, Zawidzki pays a lot of attention to the peculiar, distinctive features of human mindshaping: it generally consists in a mechanism which makes a target subject match a model (2013, p. 31) and, especially in the case of imitation, it is a quite common phenomenon in various animals; although, animals always need a *concrete* and present model, while human subjects can try to match a 'virtual', ideal model. This is quite an evidence in the law system, which takes as a model the 'ideal citizen' (2013, pp. 60-61), or widely present in the use we make of *public narratives*: "What would Jesus do in my shoes?" a Christian could think before making her choice (2013, p. 35). This is, I think, the most important point Zawidzki focuses on: in his opinion we use public narratives as 'normative systems' (like the traffic

norms in the crossroad example I gave before – see *supra*, this section). Following Victoria McGeer (2007), he suggests that these narratives mainly play a *regulative role* (Zawidzki, 2013, p. 57) among the members of a considered group. So Zawidzki, building on the concept of 'regulative ideal' coined by McGeer, mainly focuses on the mindshaping function of public narratives as a kind of *ethical*[2] regulative factor which leads each member to conform to the group narrative. Again, he argues that the main role of group narratives is to teach how someone should *act* to be compliant with his group (Zawidzki 2013, pp. 52-53, 60). This is, in my opinion, the most critical point of his view. I am going to explain why in the following section.

3. What mindshaping lacks

As I explained in the previous section, the main feature of Zawidzki's approach to mindshaping is the regulative role a model plays influencing the minds of a community. He notices that it is peculiar of humans to imitate not just actual models (as other animals do), but also 'virtual' ones: namely, *ideal models*. In this case he chooses two representative examples: the ideal citizen, presupposed by the laws system, and Jesus, an ideal model for Christians. They clearly are both *ethical models*. In fact, they both represent (within their respective frameworks) an example of 'right' behaviour; they both have an obvious *prescriptive* value. This is the reason why Zawidzki argues that public narratives play a regulative role: he thinks that narratives limit the set of cultural games (in the Wittgensteinian sense) we can play in our culture (2013, p. 58); each narrative affords just its peculiar games. Also, he maintains that narratives have a multi-level digital and sequential structure (2013, p. 58): this is, they directly represent all the steps an agent has to conform with to observe the norms of the cultural game. So, when arguing in favour of mindshaping, Zawidzki's main concern is to explain how

[2] I am using, here, the term 'ethical' in its wide etymological sense: *ethos* as custom, habit. So, when I say 'ethical models' or 'ethical behaviour', I include all the sphere of practical behaviour. This entails that I do not limit this label to mere moral models, but I include all the instances of customs transmission, such as teaching the fundamental rules of a community, but also teaching the typical strategies the community has developed, in its history, to solve some practical problems: for instance, hunting, fishing or building strategies and the like. Therefore, I use the label 'ethical mindshaping' to include also the case of apprenticeship analysed by Sterelny (2012, § 2.3).

public narratives influence our *ethical behaviour* within our group. This is because he conceives public narratives in general as abstract systems of norms which regulate our daily life, just as in the traffic norms example. Actually, the traffic norms system appears to be his prototypical model for public narratives. His perspective owes a lot to both Victoria McGeer (2007) and Kim Sterelny (2012): he claims that public narratives, like myths and laws, teach how to act through normative sanctioning (Zawidzki 2013, p. 60) (when the agent fails to abide by the norm) and represent a prototypical behaviour which is socially acceptable for all the components of a certain community. He is pretty clear on this point, when explaining that an ideal model such as the protagonist of a myth – which is an *abstraction* of all those values a community identifies itself with – is a publicly recognisable and well known character for everyone (Zawidzki 2013, p. 61). From Sterelny, he takes the idea that intergenerational learning permits a constant flow of information which brings the community rules from the elderlies to the youths. This is important to understand Zawidzki's position about the function of public narratives: Sterelny (2007, 2010, 2012) focuses on humans' ability to transmit practical knowledge from a master to an apprentice as an evolutional linchpin – which is, also, one of the most fundamental mindshaping techniques (see *supra*, VI, § 1). In this context 'practical knowledge' does not mean just hunting techniques or artisanal strategies; it is also (and more importantly) a set of rules to act fairly within the community: «When children of a culture master the narratives that it affords, what they learn are systems of self-regulation that prevail in that culture», says Zawidzki (2013, p. 58). This is perfectly coherent with McGeer's regulative ideal proposal: she argues that folk psychology is, in general, a constant attempt to understand what the others *ought to do*, according to the folk-psychological model we have in mind (normative practise) (McGeer 2007, p. 141) and, consequently, what *we* should find right to do according to that model we use to interpret our fellows (regulative practise): shortly, we feel the necessity to be coherent with our interpretative model (McGeer 2007, p. 146). This is the sense in which our interpretative frameworks have a clear regulative function in Zawidzki's view: «Our interpretations of how we and others act are simultaneously instructions for how we and others *are supposed* [my emphasis] to act, and this keeps our behavior in line

with our interpretive expectations» (2013, pp. 52-53). This is important, in his opinion, because the main reason for adjusting to a regulative model is, eventually, a reliable prediction of the behaviour of the others: if I know the normative system which is a wide accepted public narrative within our community – the traffic norms, for instance – I shall be able to predict your reaction in any possible situation resting on the normative system we share. This is the reason, I think, he focuses so much on the behavioural features of mindshaping. Also, he stresses our natural disposition to imitate the others to be more easily accepted within the group, like in the 'chameleon effect' (2013, pp. 50-53).

Although I agree with Zawidzki's general approach to mindshaping and his importance in group coordination dynamics, I also believe that contemporary discussion about mindshaping has eluded a very important issue which is the fundamental point I argue for throughout this chapter: mindshaping through images.

We have seen how much attention has been paid so far to the behavioural features of mindshaping; social *education*, roughly speaking. But this is a quite partial description of such a wide-ranging phenomenon. Mindshaping is clearly not just a form of 'chameleon effect' which should make easier for an agent to be accepted in his group; nor it is just a technique to make humans more alike through «non-conceptual, automatic mechanisms of conformism» (Zawidzki 2008, p. 204).

My point is that we do not experience mindshaping *only* when we learn the right way to seat at lunch and to eat a soup, or when we learn how to dress and speak in such and such circumstance, or when we learn to respect traffic norms or the legal system of our community, or when we learn how to fish and hunt and play football. We also do experience mindshaping when we learn how *to form concepts*, how *to represent* and how *to categorize*: we do not just passively receive mindshaping from the others (from a master to an apprentice); instead, we mindshape our fellows each time we modify our environment – especially our public narratives – to show them *our point of view*. Producing representations, we are telling them what they should do to be understandable for us. So, the regulative function of mindshaping is not just displayed at an ethical level; it also manifests at a very *cognitive* level indeed. Public narratives do not just tell us about moral examples, *exempla virtutis* which teach us the right way to

manage our life together; public narratives also *describe* how things *are*, not just how they *ought to be*. They afford us a *canon* – a rule – to represent reality. I shall call this 'categorial mindshaping'.

In the next section I shall clarify this concept and outline my proposal for a categorial mindshaping. I shall do it appealing to the role of public images in casting the narratives of a community.

4. Categorial mindshaping

Categorial mindshaping is closely related to the emergence of public narratives. As Zawidzki reminds us (see *supra*, VI, § 3), an important feature of any public narrative is its digital and sequential development: I am pretty sure he is thinking of these narratives as actual *stories*. So they are digital, because they are composed by countable moments, actions or *frames*, to borrow a metaphor from cinema; and they are sequential, because they form coherent sequences of actions. But stories are not just narrative entities, they are *descriptive* too; also, stories are not just oral (nor only written) artefacts, they are *visual* too. Listening to a story (a myth, for instance) orally shared within a community, the youths are taught by the elderlies not just about the powers of an hero or a goddess, but also about his or her very physical appearance: "White-Armed Hera", says Homer in the *Iliad*, and any ancient Greek could immediately get a concrete and defined image of how Hera would *look like* – this too being a clear example of mindshaping[3].

[3] Someone might doubt that we have here a clear example of mindshaping as I suggest, because this example might appear to some people just as a common case of activation of the imagination. However, I would respond to such possible objection with the following argument: let's consider John and James, two friends having an ordinary chat in a pub; at a certain point of the chat John suddenly says "Hey, man! Do you remember Peter? I saw him yesterday and he was wearing a very fancy scarlet jacket". James imagination would be certainly stimulated by this new piece of information, but indeed *no normative mechanism would be involved* in his imaginative process: James has no reason to start using this description John gave him as a reliable representation of Peter.
The case of Hera's description Homer gives us in the *Iliad* is, then, completely different: while John's words about Peter merely are a common description without any particular status in the cultural niche they are said, Homer is instead offering a representative description whose normative character is determined by the very position Homer's poem occupies within the ancient Greeks' cultural niche. Certainly Homer (just as Hesiod) was considered as an 'institution' and, then, all that information he

Nevertheless, any Greek could also recognize Hera or Heracles on the metopes of a temple, where their story (the public narrative) was told through a *visual* medium. And the same can be said of Christians with reference to the lives of saints, exquisitely represented in the mosaics of the marvellous Byzantine basilicas, or in the remarkable cycles of frescos of the Italian Renaissance cathedrals. But the key point here is not just the ethical mindshaping implied by the narrative level of each story: the real key point here is the categorial mindshaping entailed by its descriptive level. When you look at a metope from the cycle of Heracles, you see a *frame* of a story, but you also can learn how Heracles looked like: you do not just learn that Heracles was brave (ethical mindshaping) fighting against the centaurs, you also learn that he had long hair curled in braids and a beard (categorial mindshaping). In the same way, entering in a Byzantine basilica, you do not just learn that St. Paul was a Roman persecutor of Christians who suddenly converted to their religion; you also learn that he was almost bald and had a large beard.

So, categorial mindshaping refers to all those descriptive elements of a public narrative (especially the visual ones) which shape our categories, teaching us *how to imagine* a certain subject. Certainly, categorial mindshaping also includes any other attribute which, by description, can mould the global image we have of that subject: this is, for instance, any public mention of flavours, smells, textures and sounds. Nevertheless, in this chapter I decided to focus just on visual elements, postponing to future investigations a more global development of my thesis which will include a detailed analysis of all these descriptive components I shall not be able to examine here. This is because I think that visual attributes in public narratives manifest a more evident *normative aspect*, being based on a public visual artefact (see *infra*, VI, § 4.2) which works as a model for our mental representations and to which anyone can appeal as a reference.

I shall argue through this section that this visual information can be codified within an *icon*, available into our ecological environment. I shall argue that we build,

---

provided about gods, heroes and cosmology was accepted – in that context – as the most reliable one. This is why it is not possible to consider Homer's description of Hera as a trivial case of imaginative stimulation, while it constitutes indeed a genuine case of categorial mindshaping.

within our ecological niche, a lot of architectures and spaces which permit us to organize *systems of icons*; I shall call these spaces *imaginal spaces*. They represent a visual version of public narratives and they do have a descriptive value (e.g. "Heracles looked like this and this") but, being publicly shared, they also get a *normative value*: they establish a public *canon* for any future representation of that subject. In this way, icons we share in our ecological (and, at this point, *cultural*) niche are a clear medium of our self-regulative practices.

In the following subsections I am going to explain why icons are so important for mindshaping; what an imaginal space is; which self-regulative practices this schema implements.

4.1. The importance of *icons*

First of all, I must clarify what an icon is[4]. The first definition which is generally used to introduce a novice to the meaning of this word is: an icon is a sign which directly refers to its referent through a similarity relationship (Eco 1975/2016, p. 309.). But, what does it mean 'similarity relationship'? This has been a thorny and largely debated question in semiotics since the early criticisms to this ambiguous Peircean definition. In 1946, Charles W. Morris defined a sign as genuinely 'iconic' as far as it has the *same* properties of its *denotata* (Eco 1975/2016, p. 306). Obviously, following this literal definition, only a duplicate could be considered a proper icon and even a portrait would merely be a partially iconic sign: in fact, it just reproduces *some* of the original features of the subject, such as his colours (but, clearly, not his texture!). Umberto Eco, in his famous *A Theory of Semiotics*[5], argues that any iconic sign just reproduces some salient properties of its *denotatum* (Eco 1975/2016, § 3.5). This means that each time we produce any iconic

[4] In this section I will not distinguish between *icons* and *hypoicons*. I will make no difference between a particular artefact and the visual information encoded in it; in fact, I consider that this distinction, so typical in Peircean semiotics, is unfortunate because it keeps the door open for Platonism. Although thinking of an icon as the visual information instantiated in various hypoicons could be a useful logic abstraction for the individuation and the analysis of that information, there is no real reason for postulating any general content as separated by concrete hypoicons.

[5] To write this chapter I used as a reference the Italian edition of this book. See (Eco 1975/2016) in the Bibliography.

representation of a concrete subject, what we are actually doing is a translation from our perception to a *graphic code*; something like a similarity transformation in Euclidean geometry (Eco 1975/2016, p. 307, pp. 309-315): the schematic drawing of a hand is a graphic translation, a bijective transformation (point by point), of the perceptive schema of the hand I see into a definite graphic code. This is the representational canon and, to perceive it, any subject needs a particular *training* (Eco 1975/2016, p. 309), namely he needs to know the *transformation rule* applied to that particular instance[6].

Starting from this premise, I shall refer to iconicity as a kind of *isomorphism* between a real object and its graphic representation, regulated by a definite transformation rule.

This isomorphic relation is crucial in my view because, just as the actual object presents a certain number of perceptual constraints so it is the case of the relative icon. What I mean is that the very icon manifests a lot of physical or, better, *graphic* constraints too (see *infra*, V, § 1).

It is generally said that the most important contribution of Gestalt psychologists (Köhler 1947) consisted in putting the emphasis on the top-down constructive component of perception, namely the subject's re-organization of proximal stimuli. Instead, I suggest that this emphasis should be put on the structural constraints that form the ground of any local perceptual instance. Let's consider for instance the famous Kanizsa's images (Kanizsa 1955): both the white triangle in Fig. 6.1 and the square in Fig. 6.2 do not physically exist; they are a perceptual top-down construction based on proximal stimuli. Nevertheless, no subject could perceive a circle instead of a triangle or a square. This is because of the physical (in this case, *graphic*) structure of the object of perception: it manifests certain structural constraints that afford us *only* this or that geometrical re-organization of the proximal stimulus.

[6] For instance, in Euclidean geometry, a similarity transformation follow this rule: "Figure A' ought to have the same angles and the same ratios between distances of Figure A".

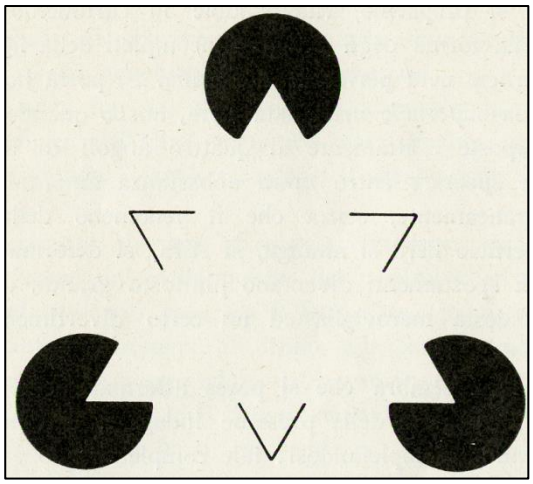

Fig. 6.3 White triangle with no-gradient borders produced by amodal completion

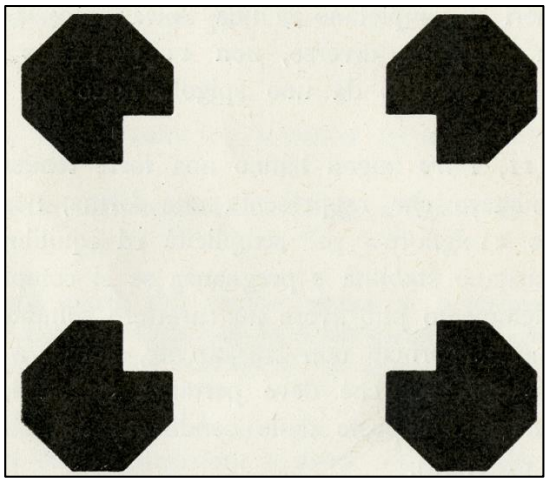

Fig. 6.2 White square with no-gradient borders produced by amodal completion

So, icons present a *structural information* which constrains our perception of their content, a content derived from the real referent of an icon by means of an isomorphism. These constraints are, then, those forcing function, those intrinsic properties Norman spoke of (Norman 1991, 2013).

Let's analyse, now, the following icon (Fig. 6.3): it represents a famous biblical episode.

Fig. 6.3 White marble low relief representing Jesus with his disciples next to the sea. Master of Cabestany

But, here is the problem: how can we decode its meaning if we have no knowledge of the transformation rule governing the isomorphism? We can do it appealing to its structural properties, the graphic constraints of the very icon. We can try to find the rule through *abduction*: Romanic sculpture is pretty symbolic and unrealistic in its representation of subjects; even though, we can focus on some unambiguous perceptual elements of this image: it is superfluous to remember here the wide literature about human specific cognitive ability in perceiving and recognising faces, even in condition of poverty of the perceptual stimulus, so it is easy to explain the immediate recognition of the three human(oid) figures carved in the marble. But all the rest of the scene is confused. What does its base represent? The unique unambiguous element is the prevalence of repeated

sinuous lines: it could be water, but also sand or even grass. Then, how can we know what kind of object is represented under the two smaller human figures? It could be a boat, a wooden pressing basin for grape, or even a striped cloth. Yet, we have an important discrimination element: there are some *animals* sketched down there (both eyes and mouths are immediately recognisable); they are fishes (Fig. 6.4).

Fig. 6.4 White marble low relief representing Jesus with his disciples next to the sea. Master of Cabestany (detail)

So, there is no doubt that the sinuous lines refer to *water*. Now the context is much more clear: those two men are in a *boat*; they are sailors or, maybe, *fishers*. So, what the graphic constraints of the icon push us to see is a scene where a man is pointing at those two fishers from the shore.

Many Christians could recognise the *isomorphic* biblical episode of Jesus calling Peter and his brother Andrew, who were fishing in the Sea of Galilee[7]. Many of them could actually recognise the figure of Jesus because of the crossed halo behind his head and the blessing gesture of his right hand (Fig. 6.5). Someone more educated could even recognise the Latin words «*Pax vobis*» (which mean «Peace to you») on the book (the Gospel) he holds in his left hand.

[7] See this biblical episode in the Gospel, Mc 1,16-20.

Fig. 6.5 White marble low relief representing Jesus with his disciples next to the sea. Master of Cabestany (detail)

Obviously, all these last symbolic elements are accessible just to those who already know the referent of this icon, namely the biblical story. But our main aim, in this section, was just to argue how the very graphic structure of an icon contains information that *shapes our minds*, forcing us to a particular interpretation of its content. This is because no graphic transformation rule is completely arbitrary: instead, the physical structure of the referent imposes its determinations as in any isomorphism[8].

In the next subsection I shall explain how a particular transformation rule acquires a normative power within a community, modifying its regulative practices and mindshaping its members.

4.2. Constructing an *imaginal* space: the normativity of canon

So now we know that an icon is connected to its referent by means of an isomorphism; also, we know that this isomorphism is regulated by a transformation rule, which is

[8] About this point, see also (Eco 1997/2016, § 2.9.)

actually the very *canon* chosen by the author of the icon to translate some relevant perceptual features of an object into its graphic representation. Yet, this isomorphism still follows a 'private' rule[9]. How could this private rule mindshape the other members of the group to which the author of the icon belongs? Apparently, we are here in the same situation of Victoria McGeer, when she tries to explain the regulative dimension of folk psychology. Her argument starts from an individual perspective too: she argues that each time that we apply a folk psychology scheme to our fellows we do not consider what they *could* do, but what they *ought to* do; then, each of us adjusts himself to his folk-psychological scheme of rules because of a coherence necessity (see *supra*, VI, § 3); this results in self-regulation. But, in spite of the apparent similarity, we cannot rely on self-regulation to explain categorial mindshaping because, in my approach, it is mediated by icons; so, my position is that each agent can mindshape his fellows redesigning the cultural niche they share. And a cultural niche is also a space characterized by a plethora of images (see *supra*, V, § 3); in this sense a cultural niche is also an imaginal space.

So my claim is that any author (namely, any icon producer) can influence the others just sharing with them his icons, which encapsulate the transformation rule, that is the canon he has chosen. Then, categorial mindshaping is possible as soon as a transformation rule is made *public*.

Nevertheless, this does not imply any normative power yet: in fact, an hypothetical agent *A* can see and immediately reject the canon afforded by the work of an agent *B*, who shares the same niche with him[10]. This is, after all, a mindshaping effect too (see *supra*, VI, § 1).

Hence, to get normative power a canon has to become *publicly recognised* as the best transformation rule. I argue that this is possible through two different salience factors: the *importance* and the *diffusion* of the icon within the considered niche. This means that a canon can be recognised as the best transformation rule by the community

[9] It is superfluous to specify that here 'private' does not mean 'internal' or 'mental'. Any transformation rule, once it is concretized into a particular icon, is in some sense already 'public'; yet, this rule could be known or applied just by the original creator of that icon and, in this sense, is 'private'.

[10] Let's think of the common case of a visitor of an art exhibition who does not understand the meaning of a painting.

if the icon is officially situated in a place[11] with particular importance and authority for the group (e.g. a temple), in such a way that it can get an institutional value[12]; or if it is massively reproduced and diffused within the cultural niche (e.g. commercial products and artisanal iconography, such as mythical figures on Greek amphorae).

Let's consider, now, Fig. 6.6: it is a metope of a Greek temple in Paestum, southern Italy; it represents Heracles defeating and killing the giant Alcyoneus.

Fig. 6.6 Heracles defeating the giant Alcyoneus

Metopes were situated in the outside of the temple, disposed in a sequential line above the columns; so, they were available for that community at any time and, by ostension (Eco 1975/2016, p. 349), they afforded the *correct canon* to represent a determinate subject. They afforded a *referent* (e.g. the figure of Alcyoneus) to people who did not have already one, or regulated the referent for those who had it. In this case the value of this icon is related to its location, but it also can be due to the diffuse presence of its instances within the niche, like in the case of the surface of ordinary products: for Greeks it could be the case of mythological figures represented on amphorae (e.g. sirens and satyrs); for us it could be the case of soft-drink bottles commercialized in our social

[11] About this point, see also (Eco 1975/2016, § 3.4.10) on 'toposensitivity' of signs.

[12] In fact, this is the reason why a personal moral disposition, belief or statement is not normative in so far as it does not belong to a legal code recognised by the community. This is because an *institutional value* is recognised in the legal code.

niche. This suggests me an interesting example: the worldly diffused icon of Santa Claus as a funny chubby old-man wearing red clothes began to spread after being reproduced on the bottles of a famous soft-drink, becoming the official canon for the representation of this character.

Then, any new icon is at first just a *proposal*[13] for a new canon. But, depending on its impact on the community, it may be recognised as the most effective representation of its referent and, hence, as a model, a *norm* for the reproduction of future instances of the same kind.

4.3. Shaping categories through images

It is clear, at this point, that by means of icons we can mindshape our fellows both in representation and categorization processes, casting for them rules for iconic reference. In this subsection I would like to focus on some salient ways to use categorial mindshaping to convey information both theoretical (e.g. information about a state of things) and practical (e.g. communicating the sense of time or a kinetic category such as 'displacement', 'fight' or 'escape').

As I explained in § 4.2, as soon as it is made public, any icon becomes *ipso facto* an act of ostension (Eco 1975/2016, p. 349). Then, it should be considered as equivalent to a *declarative proposition*. Let's consider Fig. 6.7: it is a mosaic of the cycle of the Old Testament, in the Cathedral of Monreale (Sicily); it represents the creation of stars and planets. This act of ostension is *declaring* a state of things to any spectator within the niche: it is stating "This is the Universe, these are the stars and the planets which

[13] I am using here the word 'proposal' because of the naturally ostensive feature of any icon which, as an instantiation, *offers* or *proposes* a transformation rule. Then, even though it is true that any icon is a peculiar instantiation of a determinate transformation rule, this rule should not be considered as a norm yet: in fact, despite its nomological structure, it is only a *procedural* rule (an algorithm) for the construction of a certain artefact which may (or may not) be accepted by the community. Yet, this procedural rule might become publicly recognised as a norm – a *canon* – depending on its impact within the considered cultural niche.

Fig. 6.7 Mosaic representing God's creation of stars and planets

actually exist"[14]. Hence, this icon mindshapes the agents of the niche about the 'real' structure of the Universe.

Yet, categorial mindshaping can mould kinetic categories too. This is the case of all those icons whose subject is an *action*, for instance a race or a fight, like in this fragment of a Greek frieze representing a battle between Greeks and Amazons (Fig. 6.8): the very structural information of this icon (e.g. the tension of anatomies and clothes, the position of shields and swords) is conveying the *movement of the struggle*; it affords a rule for the representation of the fight. Hence, it might shape its users to think of action and movement in terms, for instance, of muscular tension.

Fig. 6.8 Battle of Greeks and Amazons

Likewise, an icon can afford, through its graphic structure, even a canon for the representation of *time*[15]. We have a good example in Fig. 6.9, where the icon is communicating to us both the concepts of a 'before' and an 'after' by means of the

[14] Certainly, this content is context-dependent: Fig. 6.7 is just a frame of a complex sequence of scenes representing the biblical history from the beginning of time.

[15] For an introduction to the representation of time in paintings, see (Calabrese 2006, § 6).

reiteration of the same subjects both on the background, in the middle of the painting, and on the proscenium: it is the story of the Creation of Eva; then we see the Devil's temptation through the offering of the Fruit of Sin; finally, the Expulsion from the Eden. So this icon, through its structural subdivision of spaces, is displaying the *flow of time* and it is *affording a rule* for the representation of this flow: namely, it establishes a direct relation between distances in time and in space and it eventually might shape the mind of its users making them think of time in terms of space.

5. Conclusion

In this chapter I tried to outline the main reasons in favour of a mindshaping practice which, through the public use of images, moulds the very categories of our minds. I called this practice 'categorial mindshaping', as opposed to the 'ethical mindshaping' to which the majority of philosophers of folk psychology have devoted their theoretical efforts so far.

The most important commitment of this chapter is to demonstrate that even the public use of images constitutes a *normative system*, which affords representational canons to the agents of a group, eventually moulding their minds. It is clear that these norms, depending both on the structure of the icons and the one of the niche a concrete community lives within, are continuously redesigned by the agents who use them. In fact, as I highlighted at the end of the last chapter, these norms emerge accordingly to a stigmergic prototyping dynamics (see *supra*, V, § 2.2). In this sense, a community is able to produce instances of collective beliefs (see *supra*, V, § 3), that is, acts of reference in form of public representations which are collectively built within the cultural niche of a certain community and, for that, they are attributable to the whole community regarded as a collective subject. Also, public representations, once they acquire the status of canonical representations, carry out a meta-cognitive function with respect to the individual integrants of the community, whose categories are shaped in conformity with such canonical representations, so that they regulate the acts of reference acceptable for the members of a certain group.

Fig. 6.9 ***The Haywain Triptych.*** **Hieronymus Bosch (detail)**

In this chapter I also tried to demonstrate that the existence of a normative system as a consequence of the public use of icons, does not only cast 'static' categories and concepts (such as types) affording a determinate referent; instead, this use shapes individual minds giving canons to the members of a determinate community even for the representation of 'dynamic' categories such as 'time' and 'movement'. In these sense semantics, conceived as the capacity to create meanings by means of acts of reference, is an emergent property of collectives.

CONCLUSION

The emergent properties of a community

The main aim of this thesis was to defend that the very cultural structure of a certain human community, constitutes the emergent collective mind of that community itself. To argue for this thesis I drew on an *active externalism* approach.

'Active externalism' is a label proposed by Andy Clark and David Chalmers (1998). In their original perspective, it entailed a new focus on the relevance of the environment for our cognitive processes: against the 'passive' role of the environment in classic externalism (Putnam 1975), they argued for an *active role* of environmental resources in cognition. I, myself, in this thesis committed to this point.

Nevertheless, there are many ways to regard active externalism: Clark and Chalmers interpreted it as a framework to explain the relationship between human cognition and the artefacts we humans manipulate when we undertake a cognitive activity. So they argued for the extension of the mind into the artefacts we use in our cognitive routines and, then, they argued for a *constitutive* role of the external artefacts in cognition (see *supra*, I, § 1). I endorsed, instead, the complementarity principle, which defends that environmental resources do not constitute individual minds, but *complement* them. I considered, in fact, that cognitive artefacts are external *supports* for pieces of information and, then, they enable the externalization of such information into the surrounding environment, into the ecological niche where the agents live.

Also, I committed to the Third-Wave social focus on cognition and I endorsed Gallagher's definition of cognition as a meaning-making activity. I remarked, then, that it is just this meaning-making activity what distinguishes a mere instance of intelligence from a case where we have a genuine minded entity. Applying this criterion to collective entities, I noticed that there are two strategies to speak of collective minds: a first strategy that I named 'Cartesian' view, applies a psychological focus based on the model of the personal psychology in human agents; in this sense, the supporters of such a view ground their arguments on individual mental features like beliefs, desires and other

intentional states to argue in favour of the existence of similar intentional states attributable to a collective subject (see *supra*, II, § 1.2.1). By contrast, I labelled the second strategy 'informational' view, because the theorists I mentioned in this group defend that collective entities are genuinely minded when they feature a structure of information collectively manipulated within a group by its integrants and novel emergent properties which are ascribable to the whole group. Representations work as such a structure of information when they are collectively built, shared and manipulated. I committed to this informational view of the collective mind and I argued that the meanings produced by the integrants of the group crystallize in such collectively shared representations.

I explained that representations entail normativity: that is, they solicit the establishment of normative practices which regulate their use by the integrants of the community. I contended that, in the ecological space a community lives, that community also builds a cultural niche which works as a cognitive scaffold (see *supra*, III, § 3): in this scaffold, the community organizes the representations which are publicly managed by its members. I defended that public representations are physically displayed onto representational artefacts (e.g. paintings) and are collectively manipulated by each integrant of the community who interacts with them: they form a system of collective memories externally stored into such artefacts (e.g. exograms). Moreover, I stressed that representations feature intrinsic properties which delimit their possible meanings: namely, formal constraints.

I also noticed that, when public representations are structured in a cognitive scaffold, they create public narratives (see *supra*, IV, § 4) which enable the collaborative construction of collective memories into the ecological space where the community lives: in fact, narratives have a spatial dimension, they expand in the cultural niche and embrace the whole community.

I endorsed in this work a notion of belief regarded as an act of reference. I then distinguished between individual acts of reference and collective acts of reference. I argued that icons represent a particular case of act of reference. I claimed also that, in the ecological niche the community inhabits, publicly shared iconic representations may

stimulate the stigmergic emergence of a representational canon (see *supra*, V, §§ 2-3). Canonical representations, then, are to be considered acts of reference attributable to the community as a collective subject; they constitute instances of collective beliefs.

I finally argued that canonical public representations carry out a double function: at the systemic level of the community, where the collective mind emerges, they play a cognitive role because they are collective beliefs; at the individual level of the integrants of the community, instead, they play a meta-cognitive role in the sense that they regulate individual cognitive practices, moulding the minds of individual agents by means of what I called a *categorial mindshaping*.

I concluded this work with a final chapter in which, drawing on a case study, I explained how an icon may force an agent to use it just for a limited range of acts of reference, because of its intrinsic properties (namely, its formal constraints). I also further developed this point to offer an hypothesis to explain why a canon emerges as such and then becomes a normative constraint which moulds the very categories individual agents apply. Canons, then, are collectively built elements which regulate individual acts of reference, they mould the very referents the integrants of a certain community refer to and, ultimately, they mould the meanings these individuals handle. So, in addition to collective memories and collective beliefs, the very semantics, regarded as the outcome of the meaning-making activity of the group, turns out to be an emergent property of a community.

CONCLUSIÓN

Las propiedades emergentes de una comunidad

El principal objetivo de esta tesis ha consistido en defender que la propia estructura cultural de una cierta comunidad humana, constituye la mente colectiva emergente de esa misma comunidad. Para defender esta tesis he adoptado un enfoque de *externismo activo*.

'Externismo activo' es un término propuesto por Andy Clark y David Chalmers (1998). En su perspectiva original, esta implicaba una nueva mirada sobre la relevancia del entorno para nuestros procesos cognitivos: contrariamente al rol 'pasivo' del entorno en el externismo clásico (Putnam 1975), ellos defendían un *rol activo* de los recursos del entorno en la cognición. Yo también me he comprometido, en esta tesis, con este punto.

Con todo, hay muchas maneras de mirar al externismo activo: Clark y Chalmers lo han interpretado como un marco para explicar la relación entre la cognición humana y los artefactos que los humanos manipulamos cuando emprendemos una actividad cognitiva. Por ende, argumentaron en favor de la extensión de la mente en los artefactos que usamos en nuestras rutinas cognitivas y, luego, defendieron un rol *constitutivo* de los artefactos externos en la cognición (véase *supra*, I, § 1). Yo he apoyado, al contrario, el principio de complementariedad, que defiende que los recursos del entorno no constituyen las mentes individuales sino las *complementan*. He considerado, en efecto, que los artefactos cognitivos son *soportes* externos para cierta información y, por lo tanto, habilitan la externalización de dicha información en el entorno, en el nicho ecológico donde viven los agentes.

Además, me he comprometido con el enfoque social sobre la cognición propuesto por la Tercera Ola y he apoyado la definición de Gallagher de la cognición como producción de significado. He remarcado, por lo tanto, que es justo esta producción de significado lo que distingue a un mero caso de inteligencia colectiva de un caso genuino de mente colectiva. Aplicando este criterio a las entidades colectivas, he puesto en evidencia que hay dos estrategias para hablar de mentes colectivas: una primera

estrategia que he llamado perspectiva 'Cartesiana', aplica un enfoque psicológico basado en el modelo de la psicología personal de los agentes humanos; en este sentido, los defensores de dicha perspectiva apoyan sus argumentos en características mentales individuales como las creencias, los deseos y otros estados intencionales para argumentar en favor de la existencia de estados intencionales parecidos que se puedan atribuir a un sujeto colectivo (véase *supra*, II, § 1.2.1). Al contrario, he denominado pespectiva 'informacional' a la segunda estrategia, ya que los teóricos que he mencionado dentro de este grupo defienden que las entidades colectivas tienen una mente genuina cuando están caracterizadas por una estructura de información manipulada colectivamente dentro del grupo por sus integrantes y también manifiestan nuevas propiedades emergentes que se pueden adscribir al grupo entero. Las representaciones funcionan como dicha estructura de información cuando son colectivamente construidas, compartidas y manipuladas. En mi argumentación, me he comprometido con la perspectiva informacional de la mente colectiva y he defendido que los significados producidos por los integrantes del grupo cristalizan en dichas representaciones colectivamente compartidas.

He explicado que aquellas representaciones implican normatividad: es decir, estimulan el establecimiento de prácticas normativas que regulan su uso por parte de los integrantes de la comunidad. He defendido que, en el espacio ecológico donde vive una comunidad, aquella comunidad también construye un nicho cultural que funciona como un andamiaje cognitivo (véase *supra*, III, § 3): en este andamiaje, la comunidad organiza las representaciones públicamente manejadas por sus integrantes. He argumentado que las representaciones públicas están exhibidas por medio de artefactos representacionales (e.g. pinturas) y son colectivamente manipuladas por cada componente de la comunidad que interactúe con ellas: estas mismas forman un sistema de memorias colectivas almacenadas en el exterior en dichos artefactos (e.g. exogramas). Además, he remarcado que las representaciones están caracterizadas por propiedades intrínsecas que delimitan sus posibles significados, es decir, constricciones formales.

También he evidenciado que, cuando las representaciones públicas están estructuradas en un andamiaje cognitivo, crean narrativas públicas (véase *supra*, IV, § 4) que habilitan la construcción de memorias colectivas dentro del espacio ecológico donde

vive la comunidad: en efecto, las narrativas tienen una dimensión espacial, se extienden en el nicho cultural e involucran a toda la comunidad.

He sostenido en esta investigación una noción de creencia entendida como un acto de referencia. Luego, he distinguido entre actos de referencia individuales y actos de referencia colectivos. He argumentado que los iconos representan un caso particular de acto de referencia. He sostenido también que, en el nicho ecológico habitado por la comunidad, las representaciones icónicas compartidas pueden estimular la emergencia estigmérgica de un canon representacional (véase *supra*, V, §§ 2-3). Por ende, las representaciones canónicas deben de considerarse actos de referencia atribuibles a la comunidad en cuanto sujeto colectivo; estas constituyen casos de creencias colectivas.

Finalmente, he argumentado que las representaciones canónicas públicas desempeñan una función doble: a nivel sistémico de la comunidad, donde emerge la mente colectiva, juegan un papel cognitivo ya que son creencias colectivas; a nivel individual de los integrantes de la comunidad, al revés, juegan un papel meta-cognitivo en el sentido de que regulan las prácticas cognitivas individuales, moldeando las mentes de los agentes individuales por medio de lo que he definido como *categorial mindshaping*.

He concluido esta tesis con un capítulo final en el que, basándome en un caso de estudio, he explicado cómo un icono puede obligar un agente a usarlo sólo para un limitado abanico de actos de referencia, por sus propiedades intrínsecas (es decir, sus constricciones formales). También he desarrollado ulteriormente este punto para ofrecer una hipótesis para explicar por qué un canon emerge como tal y luego se vuelve una constricción normativa que moldea las propias categorías que los agentes individuales utilizan. Los cánones, por lo tanto, son elementos colectivamente construidos que regulan los actos de referencia individuales, moldean los propios referentes a los que se refieren los componentes de una comunidad y, por último, moldean los significados que estos individuos manejan. Por esto, además de memorias y creencias colectivas, la propia semántica, considerada como el resultado de una actividad de producción de significado, se revela como una propiedad emergente de una comunidad.

BIBLIOGRAPHY

Adams, Frederick, and Kenneth Aizawa. 2001. "The Bounds of Cognition." *Philosophical Psychology* 14 (1): 43–64. https://doi.org/10.1080/09515080120033571.

———. 2008. *The Bounds of Cognition*. Oxford: Wiley-Blackwell.

———. 2010. "Defending the Bounds of Cognition." In *The Extended Mind*, edited by Richard Menary, 67–80. Cambridge MA: MIT Press.

Aiello, Leslie C., and Jonathan C. K. Wells. 2002. "Energetics and the Evolution of the Genus *Homo*." *Annual Review of Anthropology* 31 (1): 323–38. https://doi.org/10.1146/annurev.anthro.31.040402.085403.

Aiello, Leslie C., and Peter Wheeler. 1995. "The Expensive-Tissue Hypothesis: The Brain and the Digestive System in Human and Primate Evolution." *Current Anthropology* 36 (2): 199–221. https://doi.org/10.1086/204350.

Ayala, Saray. 2016. "Speech Affordances: A Structural Take on How Much We Can Do with Our Words." *European Journal of Philosophy*. https://doi.org/10.1111/ejop.12186.

Bekkering, Harold. 2002. "Imitation: Common Mechanisms in the Observation and Execution of Finger and Mouth Movements." In *The Imitative Mind: Development, Evolution and Brain Bases*, edited by Andrew N. Meltzoff and Wolfgang Prinz, 163–82. Cambridge: Cambridge University Press.

Bietti, Lucas M. 2010. "Sharing Memories, Family Conversation and Interaction." *Discourse & Society* 21 (5): 499–523. https://doi.org/10.1177/0957926510373973.

Bietti, Lucas M, and John Sutton. 2015. "Interacting to Remember at Multiple Timescales: Coordination, Collaboration, Cooperation and Culture in Joint Remembering." *Interaction StudiesInteraction Studies. Social Behaviour and Communication in Biological and Artificial Systems* 16 (3): 419–50. https://doi.org/10.1075/is.16.3.04bie.

Bolici, Francesco, James Howison, and Kevin Crowston. 2016. "Stigmergic Coordination in FLOSS Development Teams: Integrating Explicit and Implicit Mechanisms." *Cognitive Systems Research* 38: 14–22. https://doi.org/10.1016/j.cogsys.2015.12.003.

Bonabeau, Eric, Guy Theraulaz, Jean-Louis Deneubourg, Serge Aron, and Scott Camazine. 1997. "Self-Organization in Social Insects." *Trends in Ecology & Evolution* 12 (5): 188–93. https://doi.org/10.1016/S0169-5347(97)01048-3.

Bonabeau, Eric, Guy Theraulaz, Jean-Louis Deneubourg, Nigel R. Franks, Oliver Rafelsberger, J. Joly, and Stéphane Blanco. 1998. "A Model for the Emergence of Pillars, Walls and Royal Chambers in Termite Nests." *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 353 (1375): 1561–76. https://doi.org/10.1098/rstb.1998.0310.

Brent, Lauren J. N., Sarah R. Heilbronner, Julie E. Horvath, Janis Gonzalez-Martinez, Angelina Ruiz-Lambides, Athy G. Robinson, J. H. Pate Skene, and Michael L. Platt. 2013. "Genetic Origins of Social Networks in Rhesus Macaques." *Scientific Reports* 3 (1): 1042. https://doi.org/10.1038/srep01042.

Brock, William A., Michael J. O'Brien, and R. Alexander Bentley. 2016. "Validating Niche-Construction Theory through Path Analysis." *Archaeological and Anthropological Sciences* 8 (4): 819–37. https://doi.org/10.1007/s12520-015-0257-0.

Bruineberg, Jelle, Anthony Chemero, and Erik Rietveld. 2018. "General Ecological Information Supports Engagement with Affordances for 'Higher' Cognition." *Synthese*, 5231–51. https://doi.org/10.1007/s11229-018-1716-9.

Calabrese, Omar. 2006. *Come Si Legge Un'opera d'arte*. Milano: Mondadori Università.

Cash, Mason. 2013. "Cognition without Borders : ' Third Wave ' Socially Distributed Cognition and Relational Autonomy." *Cognitive Systems Research* 25–26: 61–71. https://doi.org/10.1016/j.cogsys.2013.03.007.

Cavalli-Sforza, Luigi Luca. 2004. *L'evoluzione Della Cultura*. Translated by Xavier González Rovira. Turin: Codice Edizioni.

Cavalli-Sforza, Luigi Luca, and Marcus William Feldman. 1973a. "Models for Cultural Inheritance I. Group Mean and within Group Variation." *Theoretical Population Biology* 4 (1): 42–55. https://doi.org/10.1016/0040-5809(73)90005-1.

———. 1973b. "Cultural versus Biological Inheritance: Phenotypic Transmission from Parents to Children. (A Theory of the Effect of Parental Phenotypes on Children's Phenotypes)." *American Journal of Human Genetics* 25 (6): 618–37.

http://www.ncbi.nlm.nih.gov/pubmed/4797967.

———. 1983. "Paradox of the Evolution of Communication and of Social Interactivity." *Proceedings of the National Academy of Sciences* 80 (7): 2017–21. https://doi.org/10.1073/pnas.80.7.2017.

Celli, Giorgio. 2008. *La Mente Dell'ape*. Bologna: Editrice Compositori.

Chemero, Anthony. 2009. *Radical Embodied Cognitive Science*. Cambridge MA: MIT Press. https://doi.org/10.7551/mitpress/8367.001.0001.

Christensen, Lars Rune. 2013. "Stigmergy in Human Practice: Coordination in Construction Work." *Cognitive Systems Research* 21: 40–51. https://doi.org/10.1016/j.cogsys.2012.06.004.

Clark, Andy. 1997. *Being There: Putting Brain, Body, and World Together Again*. Cambridge MA: MIT Press.

———. 1998. "Magic Words: How Language Augments Human Computation." In *Language and Thought*, 162–83. Cambridge University Press. https://doi.org/10.1017/CBO9780511597909.011.

———. 2003. *Natural-Born Cyborgs*. Oxford: Oxford University Press.

———. 2010a. "Coupling, Constitution, and the Cognitive Kind: A Reply to Adams and Aizawa." In *The Extended Mind*, edited by Richard Menary, 81–99. Cambridge MA: MIT Press.

———. 2010b. "Memento 's Revenge: The Extended Mind, Extended." In *The Extended Mind*, edited by Richard Menary, 43–66. Cambridge MA: MIT Press.

Clark, Andy, and David Chalmers. 1998. "The Extended Mind." *Analysis* 58 (1): 7–19. https://www.jstor.org/stable/3328150.

Cole, Michael. 1998. *Cultural Psychology: A Once and Future Discipline*. Cambridge MA: Harvard University Press.

Cole, Michael, and James V. Wertsch. 1996. "Beyond the Individual-Social Antinomy in Discussions of Piaget and Vygotsky." *Human Development* 39 (5): 250–56. https://doi.org/10.1159/000278475.

Dennett, Daniel. 1991. *Consciousness Explained*. Boston: Little, Brown and Co.

Donald, Merlin. 1991. *Origins of the Modern Mind: Three Stages in the Evolution of*

Culture and Cognition. Cambridge MA: Harvard University Press. https://doi.org/10.1017/S0140525X00032647.

———. 1993. "Human Cognitive Evolution: What We Were, What We Are Becoming." *Social Research* 60 (1): 143–70.

———. 1997. "The Mind Considered from a Historical Perspective: Human Cognitive Phylogenesis and the Possibility of Continuing Cognitive Evolution." In *The Future of the Cognitive Revolution*, edited by David Martel Johnson and Christina E. Erneling, 355–65. Oxford: Oxford University Press.

———. 1998a. "Hominid Enculturation and Cognitive Evolution." In *Cognition and Material Culture: The Archaeology of Symbolic Storage*, edited by Colin Renfrew and Chris Scarre, 7–17. Cambridge: McDonald Institute for Archaeological Research.

———. 1998b. "Material Culture and Cognition: Concluding Thoughts." In *Cognition and Material Culture: The Archaeology of Symbolic Storage*, edited by Colin Renfrew and Chris Scarre, 181–87. Cambridge: McDonald Institute for Archaeological Research.

———. 2000. "The Central Role of Culture in Cognitive Evolution: A Reflection on the Myth of the 'Isolated Mind.'" In *Culture, Thought, and Development*, edited by Larry Nucci, Geoffrey B. Saxe, and Elliot Turiel, 19–38. New York: Psychology Press. https://doi.org/https://doi.org/10.4324/9781410603470.

———. 2001a. "Il Senso Del Sé: L'ordito Della Cultura, La Trama Della Mente." In *Le Tattiche Dei Sensi*, edited by Brunella Antomarini, 175. Roma: Manifestolibri.

———. 2001b. "Memory Palaces: The Revolutionary Function of Libraries." *Queen's Quarterly*.

———. 2007. "Evolutionary Origins of the Social Brain." In *Social Brain Matters. Stances on the Neurobiology of Social Cognition*, edited by Oscar Vilarroya and Francesc Forn i Argimon, 215–22. Amsterdam - New York: Rodopi.

———. 2010. "The Exographic Revolution : Neuropsychological Sequelae." In *The Cognitive Life of Things Recasting the Boundaries of the Mind*, edited by Lambros Malafouris and Colin Renfrew, 71–80. Cambridge: McDonald Institute for Archaeological Research.

———. 2017. “Key Cognitive Preconditions for the Evolution of Language.” *Psychonomic Bulletin and Review* 24 (1): 204–8. https://doi.org/10.3758/s13423-016-1102-x.

Doyle, Margery J., and Leslie Marsh. 2013. “Stigmergy 3.0: From Ants to Economies.” *Cognitive Systems Research* 21: 1–6. https://doi.org/10.1016/j.cogsys.2012.06.001.

Dunbar, R. I. M., and Susanne Shultz. 2007. “Evolution in the Social Brain.” *Science* 317 (5843): 1344–47. https://doi.org/10.1126/science.1145463.

Eco, Umberto. 1975/2016. *Trattato Di Semiotica Generale*. Milano: La Nave di Teseo.

———. 1997/2016. *Kant e l’ornitorinco*. Milano: La Nave di Teseo.

———. 2013. *Mnemotecniche e Rebus*. San Marino: Guaraldi - San Marino University Press.

Feldman, Marcus William, and Luigi Luca Cavalli-Sforza. 1976. “Cultural and Biological Evolutionary Processes, Selection for a Trait under Complex Transmission.” *Theoretical Population Biology* 9 (2): 238–59. https://doi.org/10.1016/0040-5809(76)90047-2.

———. 1984. “Cultural and Biological Evolutionary Processes: Gene-Culture Disequilibrium.” *Proceedings of the National Academy of Sciences* 81 (5): 1604–7. https://doi.org/10.1073/pnas.81.5.1604.

Ferber, Jacques. 1995. *Les Systèmes Multi-Agents*. Paris: InterEditions.

Floridi, Luciano. 2009. *Infosfera : Etica e Filosofia Nell’età Dell’informazione. Infosfera : Etica e Filosofia Nell’età Dell’informazione*. Digitalica ; 8. Torino: Giappichelli.

———. 2011. *Information : A Very Short Introduction. Information : A Very Short Introduction*. Oxford: Oxford University Press.

Fodor, Jerry A. 1975. *The Language of Thought*. Cambridge MA: Harvard University Press.

Francesco, Michele Di, and Giulia Piredda. 2012. *La Mente Estesa*. Milano: Mondadori Università.

Gallagher, Shaun. 2005. *How the Body Shapes the Mind. How the Body Shapes the Mind*. New York: Oxford University Press. https://doi.org/10.1093/0199271941.001.0001.

———. 2013. “The Socially Extended Mind.” *Cognitive Systems Research* 25–26

(December): 4–12. https://doi.org/10.1016/j.cogsys.2013.03.008.

Gallagher, Shaun, and Dan Zahavi. 2008. *The Phenomenological Mind*. Translated by Patrizia Pedrini. New York: Routledge.

Gallese, Vittorio, and Corrado Sinigaglia. 2010. "The Bodily Self as Power for Action." *Neuropsychologia* 48 (3): 746–55. https://doi.org/10.1016/j.neuropsychologia.2009.09.038.

Gattis, Merideth, Harold Bekkering, and Andreas Wohlschläger. 2002. "Goal-Directed Imitation." In *The Imitative Mind : Development, Evolution and Brain Bases*, edited by Andrew N. Meltzoff and Wolfgang Prinz, 183–205. Cambridge: Cambridge University Press.

Geertz, Clifford. 1973. *The Interpretation of Cultures*. New York: Basic Books.

Gibson, James J. 1979. *The Ecological Approach to Visual Perception*.

Golonka, Sabrina. 2015. "Laws and Conventions in Language-Related Behaviors." *Ecological Psychology* 27 (3): 236–50. https://doi.org/10.1080/10407413.2015.1068654.

Gordon, Deborah M. 2015. "Collective Behavior in Animals: An Ecological Perspective." In *Handbook of Collective Intelligence*, edited by Thomas W. Malone and Michael S. Bernstein, 43–51. Cambridge MA: MIT Press.

Grassé, Pierre-P. 1959. "La Reconstruction Du Nid et Les Coordinations Interindividuelles ChezBellicositermes Natalensis EtCubitermes Sp. La Théorie de La Stigmergie: Essai d'interprétation Du Comportement Des Termites Constructeurs." *Insectes Sociaux* 6 (1): 41–80. https://doi.org/10.1007/BF02223791.

Harris, Roy. 1989. "How Does Writing Restructure Thought?" *Language & Communication* 9 (2–3): 99–106. https://doi.org/10.1016/0271-5309(89)90012-8.

Havelock, Eric A. 1986. *The Muse Lears to Write. Reflections on Orality and Literacy from Antiquity to the Present*. New Haven: Yale University Press.

Heersmink, Richard. 2013. "A Taxonomy of Cognitive Artifacts: Function, Information, and Categories." *Review of Philosophy and Psychology* 4 (3): 465–81. https://doi.org/10.1007/s13164-013-0148-1.

———. 2017. "Distributed Selves: Personal Identity and Extended Memory Systems." *Synthese* 194 (8): 3135–51. https://doi.org/10.1007/s11229-016-1102-4.

Heras-Escribano, Manuel. 2019. *The Philosophy of Affordances. The Philosophy of Affordances*. Cham (Switzerland): Palgrave Macmillan. https://doi.org/10.1007/978-3-319-98830-6.

Heras-Escribano, Manuel, and Manuel de Pinedo. 2016. "Are Affordances Normative?" *Phenomenology and the Cognitive Sciences* 15 (4): 565–89. https://doi.org/10.1007/s11097-015-9440-0.

———. 2018. "Affordances and Landscapes: Overcoming the Nature-Culture Dichotomy through Niche Construction Theory." *Frontiers in Psychology* 8 (JAN): 1–15. https://doi.org/10.3389/fpsyg.2017.02294.

Heylighen, Francis. 2016a. "Stigmergy as a Universal Coordination Mechanism I: Definition and Components." *Cognitive Systems Research* 38: 4–13. https://doi.org/10.1016/j.cogsys.2015.12.002.

———. 2016b. "Stigmergy as a Universal Coordination Mechanism II: Varieties and Evolution." *Cognitive Systems Research* 38: 50–59. https://doi.org/10.1016/j.cogsys.2015.12.007.

Heylighen, Francis, and Marta Lenartowicz. 2017. "The Global Brain as a Model of the Future Information Society: An Introduction to the Special Issue." *Technological Forecasting and Social Change* 114: 1–6. https://doi.org/10.1016/j.techfore.2016.10.063.

Holland, John H. 2006. "Studying Complex Adaptive Systems." *Journal of Systems Science and Complexity* 19 (1): 1–8. https://doi.org/10.1007/s11424-006-0001-z.

Holland, O., and C. Melhuish. 1999. "Stigmergy, Self-Organization, and Sorting in Collective Robotics." *Artificial Life V* 202: 173–202.

Hölldobler, Bert, and Edward O. Wilson. 1994. *Journey to the Ants: A Story of Scientific Exploration*. Cambridge MA: Harvard University Press.

———. 2009. *The Super-Organism: The Beauty, Elegance, and Strangeness of Insect Societies*. New York: W.W. Norton & Company.

Huebner, Bryce. 2013. "Socially Embedded Cognition." *Cognitive Systems Research* 25–26: 13–18. https://doi.org/10.1016/j.cogsys.2013.03.006.

———. 2014. *Macrocognition. A Theory of Distributed Minds and Collective*

Intentionality. New York: Oxford University Press.

Hutchins, Edwin. 1995. *Cognition in the Wild*. Cambridge MA: MIT Press.

———. 2011. "Enculturating the Supersized Mind." *Phylosophical Studies*, no. 152: 437–46. https://doi.org/10.1007/s11098-010-9599-8.

———. 2014. "The Cultural Ecosystem of Human Cognition." *Philosophical Psychology*. Taylor & Francis. https://doi.org/10.1080/09515089.2013.830548.

Ingold, Tim. 2007. "Earth, Sky, Wind, and Weather." *The Journal of the Royal Anthropological Institute* 13 (Wind, Life, Health: Anthropological and Historical Perspectives): S19–38.

———. 2012. "Toward an Ecology of Materials." *Annual Review of Anthropology* 41 (1): 427–42. https://doi.org/10.1146/annurev-anthro-081309-145920.

———. 2013. *Making : Anthropology, Archaeology, Art and Architecture*. New York: Routledge. https://ebookcentral.proquest.com/lib/ugr/detail.action?docID=1170339.

Iriki, Atsushi, Michio Tanaka, and Yoshiaki Iwamura. 1996. "Coding of Modified Body Schema during Tool Use by Macaque Postcentral Neurones." *NeuroReport* 7 (14): 2325–30.

Jeffares, Ben. 2010. "The Co-Evolution of Tools and Minds: Cognition and Material Culture in the Hominin Lineage." *Phenomenology and the Cognitive Sciences* 9 (4): 503–20. https://doi.org/10.1007/s11097-010-9176-9.

Kanizsa, Gaetano. 1955. "Margini Quasi-Percettivi in Campi Con Stimolazione Omogenea." *Rivista Di Psicologia* 49 (1): 7–30.

Kauffman, Stuart A. 2000. *Investigations*. Oxford: Oxford University Press.

Kearney, Michael, and Warren Porter. 2009. "Mechanistic Niche Modelling: Combining Physiological and Spatial Data to Predict Species' Ranges." *Ecology Letters* 12 (4): 334–50. https://doi.org/10.1111/j.1461-0248.2008.01277.x.

Keesing, Roger M. 1974. "Theories of Culture." *Annual Review of Anthropology* 3: 73–97.

Kendal, Jeremy, Jamshid J Tehrani, and F. John Odling-Smee. 2011. "Human Niche Construction in Interdisciplinary Focus." *Philosophical Transactions of the Royal Society B: Biological Sciences* 366 (1566): 785–92. https://doi.org/10.1098/rstb.2010.0306.

Kiemen, Mixel, and Pieter Ballon. 2012. "Living Labs & Stigmergic Prototyping: Towards a Convergent Approach." In *ISPIM Conference Proceedings*, 1–13. Barcelona. http://www.mixel.be/files/pdf/ISPIM2012.pdf.

Kilian, Klaus. 1988. "The Emergence of Wanax Ideology in the Mycenaean Palaces." *Oxford Journal of Archaeology* 7 (3): 291–302. https://doi.org/10.1111/j.1468-0092.1988.tb00182.x.

Kirsh, David. 1995. "The Intelligent Use of Space." *Artificial Intelligence*, no. 73: 31–68.

———. 1996. "Adapting the Environment Instead of Oneself." *Adaptive Behavior* 4 (3/4): 415–52.

Kirsh, David, and Paul Maglio. 1994. "On Distinguishing Epistemic from Pragmatic Action." *Cognitive Science*. https://doi.org/10.1016/0364-0213(94)90007-8.

Köhler, Wolfgang. 1947. *Gestalt Psychology; an Introduction to New Concepts in Modern Psychology, Rev. Ed. Gestalt Psychology; an Introduction to New Concepts in Modern Psychology, Rev. Ed.* Oxford, England: Liveright.

Krueger, Joel. 2013. "Ontogenesis of the Socially Extended Mind." *Cognitive Systems Research* 25–26: 40–46. https://doi.org/10.1016/j.cogsys.2013.03.001.

Laland, Kevin Neville. 2004. "Social Learning Strategies." *Animal Learning & Behavior* 32 (1): 4–14. https://doi.org/10.3758/BF03196002.

———. 2017. *Darwin's Unfinished Symphony: How Culture Made the Human Mind*. Princeton: Princeton University Press. https://press.princeton.edu/books/hardcover/9780691151182/darwins-unfinished-symphony.

Laland, Kevin Neville, Blake Matthews, and Marcus W. Feldman. 2016. "An Introduction to Niche Construction Theory." *Evolutionary Ecology* 30 (2): 191–202. https://doi.org/10.1007/s10682-016-9821-z.

Laland, Kevin Neville, and Michael J. O'Brien. 2011. "Cultural Niche Construction: An Introduction." *Biological Theory* 6 (3): 191–202. https://doi.org/10.1007/s13752-012-0026-6.

Lashley. 1950. "In Search of the Engram." In *Society for Experimental Biology Symposium 4: Physiological Mechanisms in Animal Behavior*, 454–82. Cambridge:

Cambridge University Press.

Laskowski, Kate L, and Jonathan N Pruitt. 2014. "Evidence of Social Niche Construction : Persistent and Repeated Social Interactions Generate Stronger Personalities in a Social Spider." *Proc. R. Soc. B* 281: 20133166. https://doi.org/http://dx.doi.org/10.1098/rspb.2013.3166.

Lévy, Pierre. 1997. *L'intelligence Collective. Pour Une Anthropologie Du Cyberespace*. Paris: La Découverte.

Lewis, Ted G. 2013. "Cognitive Stigmergy: A Study of Emergence in Small-Group Social Networks." *Cognitive Systems Research* 21: 7–21. https://doi.org/10.1016/j.cogsys.2012.06.002.

Lewis, Ted G., and Leslie Marsh. 2016. "Human Stigmergy: Theoretical Developments and New Applications." *Cognitive Systems Research* 38: 1–3. https://doi.org/10.1016/j.cogsys.2015.12.001.

Lipatov, Mikhail, Melissa J Brown, and Marcus William Feldman. 2011. "The Influence of Social Niche on Cultural Niche Construction: Modelling Changes in Belief about Marriage Form in Taiwan." *Philosophical Transactions of the Royal Society B: Biological Sciences* 366 (1566): 901–17. https://doi.org/10.1098/rstb.2010.0303.

List, Christian. 2015. "What Is It like to Be a Group Agent?"

List, Christian, and Philip Pettit. 2011. *Group Agency*. New York: Oxford University Press.

Lo, Andrew W. 2015. "The Wisdom of Crowds vs. the Madness of Mobs." In *Handbook of Collective Intelligence*, edited by Thomas W. Malone and Michael S. Bernstein, 21–37. Cambridge MA: MIT Press.

Lovelock, James. 1979. *Gaia. A New Look at Life on Earth*. Oxford: Oxford University Press.

Malafouris, Lambros. 2013. *How Things Shape the Mind : A Theory of Material Engagement*. Cambridge MA: MIT Press. https://ebookcentral.proquest.com/lib/ugr/detail.action?docID=3339639.

———. 2019. "Mind and Material Engagement." *Phenomenology and the Cognitive Sciences* 18 (1): 1–17. https://doi.org/10.1007/s11097-018-9606-7.

Malone, Thomas W., and Michael S. Bernstein, eds. 2015. *Handbook of Collective Intelligence*. Cambridge MA: MIT Press.

Mameli, Matteo. 2001. "Mindreading, Mindshaping, and Evolution." *Biology and Philosophy* 16: 597–628.

Marsh, Leslie, and Christian Onof. 2008. "Stigmergic Epistemology, Stigmergic Cognition." *Cognitive Systems Research* 9 (1–2): 136–49. https://doi.org/10.1016/j.cogsys.2007.06.009.

McGeer, Victoria. 2007. "The Regulative Dimension of Folk Psychology." In *Folk Psychology Re-Assessed*, edited by Daniel D. Hutto and Matthew Ratcliffe, 137–56. Dordrecht: Springer.

Meini, Cristina. 2012. *Fuori Di Testa*. Milano: Mondadori Università.

Menary, Richard. 2007. "Writing as Thinking." *Language Sciences* 29 (5): 621–32. https://doi.org/10.1016/j.langsci.2007.01.005.

———. 2010a. "Cognitive Integration and the Extended Mind." In *The Extended Mind*, edited by Richard Menary, 227–43. Cambridge MA: MIT Press.

———. 2010b. "Dimensions of Mind." *Phenomenology and the Cognitive Sciences* 9 (4): 561–78. https://doi.org/10.1007/s11097-010-9186-7.

———. 2010c. "Introduction: The Extended Mind in Focus." In *The Extended Mind*, edited by Richard Menary, 1–25. Cambridge MA: MIT Press.

———. , ed. 2010d. *The Extended Mind*. Cambridge MA: MIT Press.

———. 2013. "Cognitive Integration , Enculturated Cognition and the Socially Extended Mind Q." *Cognitive Systems Research* 25–26: 26–34. https://doi.org/10.1016/j.cogsys.2013.05.002.

Menary, Richard, and Alexander James Gillett. 2016. "Embodying Culture: Integrated Cognitive Systems and Cultural Evolution," 1–34.

Merleau-Ponty, Maurice. 2003. *Fenomenologia Della Percezione*. Translated by Andrea Bonomi. Milano: Bompiani.

Michaelian, Kourken, and John Sutton. 2013. "Distributed Cognition and Memory Research: History and Current Directions." *Review of Philosophy and Psychology* 4 (1): 1–24. https://doi.org/10.1007/s13164-013-0131-x.

Millikan, Ruth Garrett. 1984. *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge MA: MIT Press.

Minsky, Marvin. 1988. *The Society of Mind*. New York: Simon and Schuster.

Norman, Donald A. 1991. "Cognitive Artifacts." In *Designing Interaction: Psychology at the Human-Computer Interface*, edited by John M. Carrol, 17–38. Cambridge: Cambridge University Press.

———. 1999. "Affordance, Conventions, and Design." *Interactions* 6 (3): 38–42. https://doi.org/10.1145/301153.301168.

———. 2013. *The Design of Everyday Things*. New York: Basic Books. https://doi.org/10.1002/hfm.20127.

Nowak, Martin A. 2006. "Five Rules for the Evolution of Cooperation." *Science* 314 (2006): 1560–63. https://doi.org/10.1126/science.1133755.

O'Brien, Michael J., and Kevin Neville Laland. 2012. "Genes , Culture , and Agriculture An Example of Human Niche Construction." *Current Anthropology* 53 (4): 434–70. https://doi.org/10.1086/666585.

Odling-Smee, F. John, Kevin Neville Laland, and Marcus William Feldman. 2003. *Niche Construction : The Neglected Process in Evolution*. Princeton: Princeton University Press.

Okasha, Samir. 2006. *Evolution and the Levels of Selection*. Oxford: Oxford University Press.

Omicini, Andrea, Alessandro Ricci, and Mirko Viroli. 2008. "Artifacts in the A&A Meta-Model for Multi-Agent Systems." *Autonomous Agents and Multi-Agent Systems* 17 (3): 432–56. https://doi.org/10.1007/s10458-008-9053-x.

Palermos, Spyridon Orestis. 2014. "Knowledge and Cognitive Integration." *Synthese* 191 (8): 1931–51. https://doi.org/10.1007/s11229-013-0383-0.

Parunak, H Van Dyke. 2006. "A Survey of Environments and Mechanisms for Human-Human Stigmergy." In *E4MAS: International Workshop on Environments for Multi-Agent Systems*, edited by Danny Weyns, H Van Dyke Parunak, and Fabien Michel, 163–86. Dordrecht: Springer Netherlands.

Passino, Kevin M. 2010. "Honey Bee Swarm Cognition: Decision-Making Performance

and Adaptation." *International Journal of Swarm Intelligence Research* 1 (2): 80–97.

Powers, Simon T, Jason Noble, Jordi Arranz, and Manuel De Pinedo Garcia. 2010. "Explaining Cooperative Groups via Social Niche Construction."

Prinz, Wolfgang. 2002. "Experimental Approaches to Imitation." In *The Imitative Mind : Development, Evolution and Brain Bases*, edited by Andrew N. Meltzoff and Wolfgang Prinz, 143–62. Cambridge: Cambridge University Press.

Putnam, Hilary. 1975. "The Meaning of 'Meaning.'" *Language, Mind, and Knowledge. Minnesota Studies in the Philosophy of Science* 7: 131–93.

Reed, Catherine L. 2002. "What Is the Body Schema?" In *The Imitative Mind : Development, Evolution and Brain Bases*, edited by Andrew N Meltzoff and Wolfgang Prinz, 233–43. Cambridge: Cambridge University Press.

Remotti, Francesco. 2011. *Cultura. Dalla Complessità All'impoverimento*. Roma-Bari: Laterza.

Ricci, Alessandro, Andrea Omicini, Mirko Viroli, Luca Gardelli, and Enrico Oliva. 2007. "Cognitive Stigmergy: Towards a Framework Based on Agents and Artifacts." In *Environments for Multi-Agent Systems III*, edited by Danny Weyns, H Van Dyke Parunak, and Fabien Michel, 124–40. Berlin, Heidelberg: Springer Berlin Heidelberg.

Rietveld, Erik, and Julian Kiverstein. 2014. "A Rich Landscape of Affordances." *Ecological Psychology* 26 (4): 325–52. https://doi.org/10.1080/10407413.2014.958035.

Rizzolatti, Giacomo, Luciano Fadiga, Leonardo Fogassi, and Vittorio Gallese. 2002. "From Mirror Neurons to Imitation: Facts and Speculations." In *The Imitative Mind : Development, Evolution and Brain Bases*, edited by Andrew N. Meltzoff and Wolfgang Prinz, 247–66. Cambridge: Cambridge University Press.

Rizzolatti, Giacomo, and Corrado Sinigaglia. 2006. *So Quel Che Fai. Il Cervello Che Agisce e i Neuroni Specchio*. Milano: Raffaello Cortina Editore.

Rowlands, Mark. 2010. "Consciousness, Broadly Construed." In *The Extended Mind*, edited by Richard Menary, 271–94. Cambridge MA: MIT Press.

Rupert, Robert D. 2004. "Challenges to the Hypothesis of Extended Cognition." *Journal of Philosophy*. https://doi.org/10.5840/jphil2004101826.

———. 2009. *Cognitive Systems and the Extended Mind. Cognitive Systems and the Extended Mind*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195379457.001.0001.

———. 2010a. "Extended Cognition and the Priority of Cognitive Systems." *Cognitive Systems Research*. https://doi.org/10.1016/j.cogsys.2010.04.002.

———. 2010b. "Representation in Extended Cognitive Systems: Does the Scaffolding of Language Extend the Mind?" In *The Extended Mind*, edited by Richard Menary, 325–53. Cambridge MA: MIT Press.

Ryan, Paul A., Simon T Powers, and Richard A. Watson. 2016. "Social Niche Construction and Evolutionary Transitions in Individuality." *Biology & Philosophy* 31 (1): 59–79. https://doi.org/10.1007/s10539-015-9505-z.

Searle, John R. 1995. *The Construction of Social Reality*. New York: The Free Press.

———. 2006. "Social Ontology: Some Basic Principles." *Anthropological Theory* 6 (1): 12–29. https://doi.org/10.1177/1463499606061731.

Secretan, Jimmy. 2013. "Stigmergic Dimensions of Online Creative Interaction." *Cognitive Systems Research* 21: 65–74. https://doi.org/10.1016/j.cogsys.2012.06.006.

Seeley, Thomas D., P. Kirk Visscher, and Kevin M. Passino. 2006. "Group Decision Making in Honey Bee Swarms." *American Scientist* 94 (3): 220–29. https://doi.org/10.1511/2006.3.220.

Sterelny, Kim. 2007. "Social Intelligence , Human Intelligence and Niche Construction." *Phil. Trans. R. Soc. B* 362: 719–30. https://doi.org/10.1098/rstb.2006.2006.

———. 2010. "Minds: Extended or Scaffolded?" *Phenomenology and the Cognitive Sciences* 9 (4): 465–81. https://doi.org/10.1007/s11097-010-9174-y.

———. 2011. "From Hominins to Humans: How Sapiens Became Behaviourally Modern." *Philosophical Transactions of the Royal Society B: Biological Sciences* 366 (1566): 809–22. https://doi.org/10.1098/rstb.2010.0301.

———. 2012. *The Evolved Apprentice : How Evolution Made Humans Unique. The Evolved Apprentice : How Evolution Made Humans Unique*. Cambridge MA: MIT Press.

Susi, Tarja. 2016. "Social Cognition, Artefacts, and Stigmergy Revisited: Concepts of

Coordination." *Cognitive Systems Research* 38: 41–49. https://doi.org/10.1016/j.cogsys.2015.12.006.

Sutton, John. 2006a. "Distributed Cognition Domains and Dimensions." *Pragmatics & Cognition* 14 (2): 235–247.

———. 2006b. "Introduction: Memory, Embodied Cognition, and the Extended Mind." *Philosophical Psychology* 19 (3): 281–89. https://doi.org/10.1080/09515080600702550.

———. 2010. "Exograms and Interdisciplinarity: History, the Extended Mind, and the Civilizing Process." In *The Extended Mind*, edited by Richard Menary, 189–225. Cambridge MA: MIT Press.

———. 2014. "Remembering as Public Practice : Wittgenstein , Memory , and Distributed Cognitive Ecologies." In *Mind, Language, and Action: Proceedings of the 36th Wittgenstein Symposium*, edited by V. A. Munz, D. Moyal-Sharrock, and A. Coliva, 409–44. De Gruyter.

Sutton, John, Celia B. Harris, Paul G. Keil, and Amanda J. Barnier. 2010. "The Psychology of Memory, Extended Cognition, and Socially Distributed Remembering." *Phenomenology and the Cognitive Sciences* 9 (4): 521–60. https://doi.org/10.1007/s11097-010-9182-y.

Szanto, Thomas. 2014. "How to Share a Mind: Reconsidering the Group Mind Thesis." *Phenomenology and the Cognitive Sciences* 13 (1): 99–120. https://doi.org/10.1007/s11097-013-9323-1.

Theiner, Georg. 2011. *Res Cogitans Extensa. A Philosophical Defense of the Extended Mind Thesis*. Frankfurt am Main: Peter Lang.

Theiner, Georg, Colin Allen, and Robert L. Goldstone. 2010. "Recognizing Group Cognition." *Cognitive Systems Research* 11 (4): 378–95. https://doi.org/10.1016/j.cogsys.2010.07.002.

Theraulaz, Guy, and Eric Bonabeau. 1999. "A Brief History of Stigmergy." *Artificial Life* 5 (2): 97–116. https://doi.org/10.1162/106454699568700.

Townsend Peterson, A., Jorge Soberón, Richard G. Pearson, Robert P. Anderson, Enrique Martínez-Meyer, Miguel Nakamura, and Miguel Bastos Araújo. 2011. *Ecological Niches and Geographic Distributions*. Princeton: Princeton University

Press. https://muse.jhu.edu/book/30964#info_wrap.

Trianni, Vito, Elio Tuci, Kevin M. Passino, and James A.R. Marshall. 2011. "Swarm Cognition: An Interdisciplinary Approach to the Study of Self-Organising Biological Collectives." *Swarm Intelligence* 5 (1): 3–18. https://doi.org/10.1007/s11721-010-0050-8.

Tummolini, Luca, and Cristiano Castelfranchi. 2007. "Trace Signals: The Meanings of Stigmergy." In *Environments for Multi-Agent Systems III*, edited by Danny Weyns, H Van Dyke Parunak, and Fabien Michel, 141–56. Berlin, Heidelberg: Springer Berlin Heidelberg.

Tversky, Barbara, Julie Bauer Morrison, and Jeff Zacks. 2002. "On Bodies and Events." In *The Imitative Mind : Development, Evolution and Brain Bases*, edited by Andrew N. Meltzoff and Wolfgang Prinz, 221–32. Cambridge: Cambridge University Press.

Vaccari, Andres Pablo. 2017. "Against Cognitive Artifacts: Extended Cognition and the Problem of Defining 'Artifact.'" *Phenomenology and the Cognitive Sciences* 16 (5). https://doi.org/10.1007/s11097-016-9484-9.

Vogt, Stefan. 2002. "Visuomotor Couplings in Object-Oriented and Imitative Actions." In *The Imitative Mind : Development, Evolution and Brain Bases*, edited by Andrew N. Meltzoff and Wolfgang Prinz, 206–20. Cambridge: Cambridge University Press.

Vygotsky, Lev Semenovič. 1978. *Mind in Society: The Developement of Higher Psychological Processes*. Edited by Michael Cole, Sylvia Scribner, Vera John-Steiner, and Ellen Souberman. Cambridge MA: Harvard University Press. 2018. *Il Processo Cognitivo*. Translated by Caterina Ranchetti. Turin: Bollati Boringhieri.

———. 1934. *Myšlenie i reč' Psichologičeskie isseledovanija*. Moskva-Leningrad. 2019. *Pensiero e Linguaggio*. Edited and translated by Luciano Mecacci. Roma-Bari: Laterza.

Wheeler, Michael. 2010. "In Defense of Extended Functionalism." In *The Extended Mind*, edited by Richard Menary, 245–70. Cambridge MA: MIT Press.

Williamson, Kellie, and John Sutton. 2014. "Embodied Collaboration in Small Groups." In *Brain Theory*, edited by C. T. Wolfe, 107–33. London: Palgrave Macmillan. https://doi.org/https://doi.org/10.1057/9780230369580_7.

Wilson, Edward O. 1975. *Sociobiology: The New Synthesis*. Cambridge MA: Harvard University Press.

Wilson, Robert A. 2010. "Meaning Making and the Mind of the Externalist." In *The Extended Mind*, edited by Richard Menary, 167–88. Cambridge MA: MIT Press.

Zawidzki, Tadeusz W. 2008. "The Function of Folk Psychology: Mind Reading or Mind Shaping?" *Philosophical Explorations* 11 (3): 193–210. https://doi.org/10.1080/13869790802239235.

———. 2013. *Mindshaping: A New Framework for Understanding Human Social Cognition*. Cambridge MA: MIT Press.

ICONOGRAPHIC REFERENCES

Fig. 2.1: Initial stage in a multi-agent system. Computer drawing by the author.

Fig. 2.2: Emergence of a new rule within the work-space. Computer drawing by the author.

Fig. 3.1: The four dimensions of niche construction. Computer drawing by the author.

Fig. 4.4: *The theatre of Memory*. Giulio Camillo Delminio (1550. Image in the public domain.

Fig. 4.5: *Stories of the Virgin Mary*. Duccio di Buoninsegna (1311). Museo dell'Opera Metropolitana (Siena, Italy). Image in the public domain.

Fig. 5.1: Roman Venus. Unknown author. Pompeii (I century). Image in the public domain.

Fig. 5.2: *Sleeping Venus*. Giorgione (1510). Gemäldegalerie (Dresden, Germany). Image in the public domain.

Fig. 5.3: *Venus of Urbino*. Tiziano Vecellio (1538). Galleria degli Uffizi (Florence, Italy). Image in the public domain.

Fig. 5.4: *La maja desnuda*. Francisco Goya (1800). Museo del Prado (Madrid, Spain). Image in the public domain.

Fig. 5.5: *Pauline Bonaparte as Victorious Venus*. Antonio Canova (1808). Galleria Borghese (Rome, Italy). Image in the public domain.

Fig. 5.6: *Olympia*. Édouard Manet (1863). Musée d'Orsay (Paris, France). Image in the public domain.

Fig. 6.1: White triangle with no-gradient borders produced by amodal completion. Source: Kanizsa 1955. Courtesy of Aspi - Archivio Storico della Psicologia Italiana

Fig. 6.2: White square with no-gradient borders produced by amodal completion. Source: Kanizsa 1955. Courtesy of Aspi - Archivio Storico della Psicologia Italiana

Fig. 6.3: White marble low relief representing Jesus with his disciples next to the sea. Master of Cabestany (XII century), Museu Frederic Marès (Barcelona, Spain). Museu Frederic Marès. © Photo: Guillem F-H. Courtesy of the Museu Frederic Marès.

Fig. 6.4: White marble low relief representing Jesus with his disciples next to the sea. Master of Cabestany (XII century), Museu Frederic Marès (Barcelona, Spain). Museu Frederic Marès. © Photo: Guillem F-H. Courtesy of the Museu Frederic Marès. Detail representing some fishes in the sea. Modified by the author.

Fig. 6.5: White marble low relief representing Jesus with his disciples next to the sea. Master of Cabestany (XII century), Museu Frederic Marès (Barcelona, Spain). Museu Frederic Marès. © Photo: Guillem F-H. Courtesy of the Museu Frederic Marès. Detail representing Jesus blessing and holding the Gospel. Modified by the author.

Fig. 6.6: Heracles defeating the giant Alcyoneus, temple metope of the Heraion of Paestum (VI century BC). Museo Archeologico Nazionale di Paestum (Paestum, Italy). Photo by Velvet 2011, reproduced under the Creative Common CC BY-SA 3.0 License

Fig. 6.7: Mosaic representing God's creation of stars and planets (XII century), Cathedral of Monreale (Palermo, Italy). Image in the public domain.

Fig. 6.8: Battle of Greeks and Amazons, Fragment of a frieze, Mausoleum of Halicarnassus (Bodrum, Turkey); now in the British Museum (London, United Kingdom). Photo by Marie-Lan Nguyen, 2011, reproduced under the terms of the Creative Commons CC-BY 2.5 License

Fig. 6.9: *The Haywain Triptych*, Hieronymus Bosch (1516), Museo del Prado (Madrid, Spain); detail representing the expulsion of Adam and Eva from the Eden. Source: Trivium. Art History Platform. Image in the public domain.

Esta tesis doctoral ha contado con la ayuda para Formación de Personal Investigador (FPI BES-2016-077237) vinculada al proyecto "Habla interna, metacognición y la concepción narrativa de la identidad" (FFI2015-65953-P), financiado por el Ministerio de Economía y Competitividad con fondos de la Unión Europea.