The Varieties of Situated Cognitive Systems

Embodied Agents, Cognitive Artifacts, and Scientific Practice

Jan Richard Heersmink, BSc., MA.

Macquarie University

Department of Cognitive Science

July 2014

Table of Content

Summary

The main goal of this thesis is to better understand the variety of situated cognitive systems consisting of embodied agents and cognitive artifacts, and to conceptualize how such artifacts and their users are integrated into systems that perform cognitive tasks (in scientific practice). To this end, I start by identifying and classifying the components of situated cognitive systems, including those that are human-made or artificial, natural, and social (chapter 2). Next, I focus on the artifactual element in cognitive systems by characterizing this class of artifacts as a functional kind, i.e., a kind of artifact that is defined purely by its function. This functional kind includes artifacts with proper and systems functions. Those with proper functions have a history of cultural selection, whereas those with system functions are improvised uses of initially non-cognitive artifacts. By drawing on artifact categorization in archaeology, I then develop a detailed taxonomy in which cognitive artifacts with similar informational properties are grouped into categories (chapters 3 & 4). Having developed this taxonomy, I present a multidimensional framework to conceptualize how embodied agents and cognitive artifacts are integrated into situated systems. This framework consists of the following dimensions: epistemic action and information flow, speed of information flow, reliability, durability, trust, procedural transparency, informational transparency, individualization, and transformation. These dimensions are all matters of degree and jointly constitute a multidimensional space in which situated cognitive systems, including those that are extended or distributed, can be located (chapter 5). I end this thesis by applying the taxonomy and multidimensional framework to first classify cognitive artifacts in molecular biology laboratories, and then to conceptualize how some of these cognitive artifacts and their users are integrated into situated systems (chapter 6).

Statement of Candidate

I certify that the work in this thesis entitled "The Varieties of Situated Cognitive Systems: Embodied Agents, Cognitive Artifacts, and Scientific Practice" has not previously been submitted for a degree, nor has it been submitted as part of requirements for a degree to any other university or institution other than Macquarie University.

I also certify that the thesis is an original piece of research and it has been written by me. Any help and assistance that I have received in my research work and the preparation of this thesis itself have been appropriately acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

.....

Jan Richard Heersmink
Sydney, July 2014
41949102

Acknowledgements

A project on the relationship between human thought and the material and social environment in which it is unfolding and situated can only start with thanking that environment. In my case, that has been quite an environment and there have been countless people who have contributed to the realization of this thesis, cognitively or otherwise. First and foremost, I wish to express my deepest gratitude to John Sutton for being an excellent supervisor and mentor. John's ideas, suggestions, guidance, and motivating words have been tremendously important and inspiring for writing this thesis. As a supervisor, or better, as a person, he is very involved and sincerely concerned with the research and progress of his students. I appreciate that a lot. It was a great pleasure and honour to be one of his students.

Secondly, I want to express my gratitude to my associate supervisor Richard Menary. My meetings with Richard and John were at times puzzling and confusing, but always stimulating. In either case, in retrospect I can say that those were the moments when I learned the most and provided me with ample food for thought. I also appreciated that Richard gave me the opportunity to co-lecture and tutor his courses "Philosophy of Cognitive Science" and "Biology, Mind, and Culture". Those courses had my name written all over them and it was a pleasure to help teach them. This thesis would have been impossible to write without John and Richard's excellent supervision.

In addition to my supervisors, various people have helped me finish this project. Philip Brey, Beth Preston, Tarja Knuuttila, Chris Sinha, and Sven Walter have provided me with literature on artifacts, functions, models, or situated cognition. Johnny Hartz Søraker invited me to give a talk at the philosophy department of the University of Twente and Sadjad Soltanzadeh invited me to give a talk at the Centre for Applied Ethics and Public Philosophy in Canberra. Peter Woelert and Sadjad commented on some of my chapters, papers, or ideas in other formats. Their help is much appreciated and has, in one way or another, shaped my ideas and this thesis.

Thanks to Albert Atkin and Jennifer Duke-Yonge for giving me the opportunity to tutor for "Critical Thinking" and to Jeanette Kennett for giving me the opportunity to lecture

and tutor for "Philosophy of Science" and "Mind, Meaning and Metaphysics". Teaching those courses was not only great fun but also made me a more analytical and critical thinker.

Credit should also go to my colleagues and fellow PhD students with an interest in (situated) cognition at Macquarie University: Amanda Barnier, Nicolas Bullot, Glenn Carruthers, Mirko Farina, Celia Harris, Michael Kirchhoff, Will Newsome, Kellie Williamson, and also to those with other philosophical or scientific interests: Nobu Akagi, Lise Marie Anderson, Nora Fieder, Tereza Hendl, Onni Hirvonen, Lars Marstaller, Chris McCarroll, Vincenzo Moscati, Monica Ricci, Melanie Rosen, Robert Ross, Anke Snoek, Jordan Taylor, Marina Trakas, Nicole Vincent, and all the others I forgot to thank. The members of the "Embodied Cognition", "Distributed Cognition", "Cognitive Anthropology", and "MacNappers" reading groups also deserve gratitude. Macquarie University is a great place to do philosophy.

My amazing wife Marije Nieuwenhuis deserves the most sincere gratitude, appreciation, and love. As newlyweds we moved to the other side of the world with just a backpack on our shoulders. So our four-year visit to Sydney was a kind of extended honeymoon, which has been overwhelming, strange, hot, surprising, and much more. It was without doubt the most challenging and rewarding period in my life. Thank you so much for always supporting me when I needed it! We have experienced many adventures together and it was absolutely fantastic to share this one with you as well. What will our next adventure be? To my parents-in-law, Tijs and Lotty Nieuwenhuis, I want to say thanks for making such a kind and beautiful person and thanks for allowing me to kidnap her to the other side of the world for four years.

Tenslotte wil ik graag mijn ouders, Jan en Jenny Heersmink, bedanken. Dankzij jullie morele, praktische, en financiele steun ben ik gekomen waar ik nu ben. Ik kan me jullie gezichtuitdrukking nog levendig herinneren toen ik aankondigde dat ik na twee opleidingen in de biologie, filosofie wilde gaan studeren. Dat leek me nou een goed idee. Ik kon me destijds niet aan de indruk onttrekken dat het jullie een minder goed idee leek, maar ondanks dat hebben jullie me altijd door dik en dun gesteund. Erg veel dank daarvoor. Zonder jullie steun was het me nooit gelukt. Jullie zijn geweldige ouders! Ook wil ik graag mijn broers, Albrecht en Harry, bedanken, o.a. voor financiele hulp en praktisch advies, maar ook omdat jullie altijd mijn grote broers blijven.

About this Thesis

Some of the work in this thesis has been published or accepted for publication. A list of publications is provided below. Revised versions of sections 2.2.3 and 3.2 of chapter 2 have been published in Heersmink (2013b) and revised versions of section 3.4 of chapter 2 have been published in Heersmink (2011; 2012a; 2013a). Revised versions of section 2 of chapter 3 and section 3.4 of chapter 4 have been published in Heersmink (2014b). A substantially revised version of chapter 4 has been published in Heersmink (2013b). A preliminary version of chapter 5 has been published in Heersmink (2012b) and a revised version of chapter 5 has been published in Heersmink (2014a).

Heersmink, R. (2011). Epistemological and Phenomenological Issues in the Use of Brain-Computer Interfaces. In C. Ess & R. Hagengruber (eds.), *Proceedings of the International Association for Computing and Philosophy* (pp. 98-102).

Heersmink, R. (2012a). Defending Extension Theory: A Response to Kiran and Verbeek. *Philosophy and Technology,* 25(1), 121-128.

Heersmink, R. (2012b). Mind and Artifact: A Multidimensional Matrix for Exploring Cognition-Artifact Relations. In J.M. Bishop & Y.J. Erden (eds.), *Proceedings of the 5th AISB Symposium on Computing and Philosophy* (pp. 54-61).

Heersmink, R. (2013a). Embodied Tools, Cognitive Tools, and Brain-Computer Interfaces. *Neuroethics,* 6(1), 207-219.

Heersmink, R. (2013b). A Taxonomy of Cognitive Artifacts: Function, Information, and Categories. *Review of Philosophy and Psychology*, 4(3), 465-481.

Heersmink, R. (2014a). Dimensions of Integration in Embedded and Extended Cognitve Systems. *Phenomenology and the Cognitive Sciences*. DOI: 10.1007/s11097-014-9355-1.

Heersmink, R. (2014b). The Metaphysics of Cognitive Artifacts. *Philosophical Explorations*. DOI: 10.1080/13869795.2014.910310.

1

Embodied Agents and Cognitive Artifacts

1. Introduction

In the last twenty years or so, there has been a shift in the cognitive sciences away from focusing on cognitive processes in the brain and towards focusing on cognitive processes involving brain, body, and environment. In *The Cambridge Handbook of Situated Cognition,* Philip Robbins and Murat Aydede (2009b) identify three distinct but related theses that characterize the situated cognition movement. First, the embodied cognition thesis, which claims that cognition sometimes depends on and is sometimes constituted by the body (e.g. Anderson 2003; Gallagher 2005). Second, the embedded cognition thesis, which claims that we sometimes delegate and offload cognitive and information-processing tasks onto the artifactual environment (e.g. Kirsh & Maglio 1994). Third, the extended and distributed cognition theses, which claim that cognitive processes, under certain conditions, are distributed across embodied agents and cognitive artifacts or other external resources (Hutchins 1995a; Clark & Chalmers 1998).

Situated cognition can thus be seen as the genus and embodied, embedded, and extended/ distributed cognition as its species. These approaches have conceptual and metaphysical consequences, since they move beyond an individualist form of cognitivism and towards a picture that involves brain, body, and environment. As a result, these approaches also have methodological consequences, because rather than merely focussing on cognitive processes in the brain, they advocate that we should

focus on the cognitive relation between the brain, body, and environment, both on a conceptual and empirical level. To better understand this relation, Andy Clark has argued that we need "a new kind of cognitive scientific collaboration involving neuroscience, physiology, and social, cultural, and technological studies in about equal measure" (Clark 2001, p. 151).

This thesis builds on Sutton's (2002, 2010) call to implement Clark's proposal by focussing mainly on situated cognitive systems consisting of embodied agents and cognitive artifacts. Embodied agents like us have limited information-storage and information-processing capacities. In order to complement these limitations, we organize our environment, use maps, diagrams, models, diaries, timetables, textbooks, calculators, computer systems, and many other artifacts to help us perform our cognitive tasks such as planning, navigating, calculating, learning, and remembering. Organizing our environment and using external artifacts to perform cognitive tasks transforms our embodied brains and gives us clear epistemic benefits, as it makes such tasks easier, faster, more reliable, or possible in the first place. In Donald Norman's (1993) words, these are the "things that make us smart" and without them our cognitive capacities would be radically different. Likewise, anthropologist Edwin Hutchins (1995a) argued that we should study cognitive activities "in the wild", i.e., as they unfold in and with their material and social environment. A complete understanding of our cognition should thus take into account the environmental structures and artifacts we use to help us perform our cognitive tasks, thereby enlarging the unit of analysis for the cognitive sciences. What exactly this larger unit of analysis is, and how it needs to be studied and conceptualized is the topic of this thesis.

2. Research Motivations

Situated cognition theory has widely studied the epistemic roles of artifacts in performing cognitive tasks, both conceptually and empirically (Donald 1991; Norman 1993; Kirsh & Maglio 1994; Hutchins 1995a; Kirsh 1995, 2009; Zhang 1997; Clark & Chalmers 1998; Clark 2008b; Wilson 2004; Rogers 2004; Menary 2007a, 2010a; Walter 2013). These theorists collectively focus on a variety of examples, including checklists, the rotation of zoids in playing Tetris, maps, gyrocompasses, radars, notebooks, post-it notes, diagrams, word-processers, sketchbooks, calculators, computer interfaces, and other artifacts. Their examples are well-chosen and show that particular artifacts play

particular roles in performing cognitive tasks. However, what is missing in the literature is an overarching framework in which to understand and classify the functional and informational properties of cognitive artifacts.

There have been some brief attempts to classify cognitive artifacts (Norman 1993; Nersessian et al 2003; Sterelny 2004; Brey 2005; Donald 2010). But, these classifications are somewhat limited as they focus exclusively on representational artifacts, which are artifacts exhibiting representational properties such as maps, calendars, thermometers, colour charts, diagrams, scientific models, and so on. We also use non-representational artifacts to aid our cognitive tasks, such as, for instance, consistently leaving one's car keys at a certain spot so that one does not have to remember where they are, or organizing one's workspace such that the location of the objects facilitate the task one is doing. Cooks, for instance, organize their utensils and ingredients so as to reduce the load on perceptual and memory processes in making a meal (Kirsh 1995). Current classifications thus give an incomplete picture of the diversity and variety of artifactual components of situated cognitive systems. Moreover, not only do they give an incomplete picture, their methodology for classifying representational artifacts is cognition-centered in that it takes a human agent and its cognitive processes and goals as point of departure and then classify artifacts on the basis of the cognitive process or goal to which the artifacts contribute. Such cognition-centered approaches do not pay a great deal of attention to the informational properties of cognitive artifacts.

Furthermore, it is vital to point out that it is not just artifacts or artifactual structures that complement our cognitive capacities. We also develop and learn what I refer to as "cognitive techniques" such as mnemonics to remember to order of the planets in our solar system, methods of loci to help remember certain items, or mentally visualizing and manipulating abacus beads in order to perform calculations. Artifacts and techniques are human-made and can, in that sense, be classified as artificial. Additionally, we use natural objects like stars to aid navigating, and other people to help us think and remember. Such cognitive scaffolds - whether artificial (i.e. artifacts and techniques), natural, or social - are important to study, not only because they make us more powerful and versatile thinkers, but also because they shape and transform our onboard cognitive system, both ontogenetically (De Cruz 2008; Menary 2010c; Kirchhoff 2011) and phylogenetically (Donald 1991; Sterelny 2003, 2012). Due to the

substantial variety in cognitive scaffolds, and a bias towards representational artifacts in current classifications, an overarching taxonomy providing a systematic understanding of the different kinds of scaffolds would be of great help for situated cognition theory. Clark has aptly put it as follows:

"The single most important task, it seems to me, is to better understand the range and variety of types of cognitive scaffolding, and the different ways in which non-biological scaffoldings can augment or impair performance on a task... The Holy Grail here is a taxonomy of different types of external prop, and a systematic understanding of how they help (and hinder) human performance" (Clark 2002, p. 29).

Clark's suggestion has two interrelated aspects. In order to acquire a systematic understanding of how different external scaffolds augment or impair performance on different cognitive tasks, we first need to have an understanding of the range and variety of types of such scaffolds. This strategy is sensible, however, it is not just *external* scaffolds (i.e. artifacts, natural objects, and social scaffolds) that help and hinder human performance. In order to develop a broader and more inclusive picture, we also need to take into account cognitive techniques. In other words, a first possible step to obtain Clark's "Holy Grail" consists of a taxonomy of different components of situated cognitive systems, including those that are artificial (i.e. artifacts and techniques), natural (e.g. navigating on the basis of the stars), or social (i.e. other agents) (compare Susi 2005). Given these reasons, the first step in this thesis is to classify the components of situated cognitive systems on a general level, resulting in a taxonomy that demonstrates and accounts for the diversity of situated cognitive systems.

Having this general and high-level taxonomy of components of situated cognitive systems in place, I continue by narrowing in on the artifactual element in situated cognitive systems. Why artifacts? Artifacts are important to further conceptualize as they are quite ubiquitous and we heavily depend on them for performing many of our cognitive tasks, perhaps more so than on other components. The metaphysical aspects of artifacts such as their function and their informational properties have not received significant attention in situated cognition theory. Perhaps this is so because the explanatory targets of situated cognition theory are situated cognitive *systems*, not *artifacts*. Whilst this explanatory focus is understandable, a better understanding of the

metaphysical properties of artifacts also allows us to better understand the overall system and would thus be beneficial for situated cognition theory. This is so because a better understanding of the functional and informational properties of artifacts provides a better grip on how these properties are utilized and integrated with the cognitive systems of their users. For example, one species of cognitive artifacts that are identified in the taxonomy are indices such as thermometers, speed meters, and barometers. These are artifacts that are directly connected to their target system. The user of such artifacts cannot change the informational content of the artifact and, for this reason, information flow between user and artifact can only be one-way. This one-way information flow limits the degree of integration between user and artifact. In contrast, the informational content of other species of cognitive artifacts such as icons and symbols can be altered by the user and, for this reaons, allows two-way or reciprocal information flow. This allows the artifact to be integrated much deeper with the cognitive system of its user.

In order to achieve a better understanding of these artifacts and the systems of which they are part, it is helpful to look at their metaphysical properties. To this end, I draw on work in analytic philosophy of technology, which has recently begun to address metaphysical properties of artifacts. However, these theorists focus on technical artifacts exhibiting pragmatic functions such as cars, screwdrivers, and hammers (e.g. Lawson 2008; Thomasson 2009; Houkes and Vermaas 2010; Kroes 2012) and have little, if anything, particular to say about *cognitive* artifacts. Analytic philosophy of technology would benefit from casting its net wider by including more kinds of artifacts in its explanandum, i.e., it would benefit by including artifacts exhibiting not only pragmatic functions but also cognitive functions.

So the situation seems to be this: we lack metaphysical knowledge about cognitive artifacts that would benefit both situated cognition theory and analytic philosophy of technology. To obtain this knowledge, I start by conceptualizing cognitive artifacts as a functional kind, i.e., a kind of artifact defined purely by its function. Building on the work of Beth Preston (1998b, 2013), I develop a pluralist notion of functional kind, one that includes artifacts with proper functions and system functions. Those with proper functions have a history of cultural selection, i.e., they are intended by their designers and selected by their users to perform a cognitive function. Examples include abacuses, calculators, computers, compasses, thermometers, radar systems, diaries, etc. Those

with system functions are improvised uses of initially non-cognitive artifacts, such as leaving a rented DVD on your desk as reminder to bring it back to the video store, using everyday objects like pencils and paperclips as stand-ins as to explain and reconstruct how an accident happened, or organizing one's workspace such that the objects one is interacting with facilitate the task one is performing.

This kind of artifact is thus functionally homogenous in that the members of this kind have as their function to aid and complement the cognitive processes of their users, but it is informationally heterogeneous in that the members of this kind exhibit different informational properties. The next step, then, is to distinguish between different categories of cognitive artifacts on the basis of the information they provide and to identify the relationship between function and information. Cognitive functions of artifacts supervene on their informational properties. A map, for instance, is used for navigating because the information it contains is helpful for navigating, a diagram is used for inference-making as the information it contains affords inference-making, abacuses are used for calculating as the information they contains is helpful for calculating, etc. Given this cognition-complementing role of external information, it is important to better understand the informational properties of artifacts. To this end, I develop a taxonomy of the artifactual elements in situated cognitive systems in which I distinguish between a number of categories. In doing so, an information-centered approach is taken, i.e., I take as my point of departure the specific informational properties of cognitive artifacts and taxonomize them on the basis of those properties. The methodology for creating this taxonomy is developed by drawing on artifact classification in archaeology. Artifact classification has not been a prominent topic in analytic philosophy of technology. Archaeology, by contrast, has long-standing and robust methods for classifying their objects of interest.

An important distinction in this taxonomy is between cognitive artifacts exhibiting representational properties and non-representational or ecological properties. Those that exhibit representational properties provide information *about* the world, i.e., they have aboutness or representational content. Examples include photographs, maps, models, compasses, shopping-lists, textbooks, radar systems, architectural blueprints, and so forth. The information that these artifacts provide has representational targets and therefore I refer to such artifacts as "representational cognitive artifacts". By contrast, those that exhibit ecological information provide information that does not

have representational content or a target. Examples include consistently leaving car keys on a certain spot in your apartment so that you know where they are, putting an article you have to read on top of the pile on your desk, leaving a book open and turned upside down so that you know where you have stopped reading, or leaving a rented DVD on your desk as a prompt to bring it back to the video store. These examples of organized environments do not exhibit aboutness or content. Rather, by putting artifacts in certain locations that are either deliberately usual or deliberately unusual, we intentionally encode information into the artifact and its location, thereby creating what I refer to as "ecological cognitive artifacts".

It is important to point out from the outset that category membership is based on the *predominant* informational property. However, a token cognitive artifact can exhibit more than one kind of information. Maps, for example, display isomorphism with their target and are thus predominantly iconic, but often they also contain language and numbers which are symbolic. So a map combines iconic and symbolic information. The taxonomy is meant to better understand and distinguish between the informational properties of the artifact. When analysing the integration between agent and artifact, all relevant informational properties should be taken into account, not only the predominant informational one.

Having developed a detailed taxonomy of cognitive artifacts, outlining a number of distinct but sometimes overlapping categories, the next step is to present the tools to conceptualize the degree to which agents and artifacts are integrated into situated systems that perform cognitive tasks. The literature on extended and distributed cognition does not pay a great deal of attention to the conditions of cognitive extension or distribution. When it does, it mainly focusses on Clark and Chalmers' (1998) conditions of trust, reliability, accessibility, and past endorsement (as well as the parity principle).

In response to this, I first argue that the parity principle, which claims that external artifacts ought to exhibit similar properties as do internal states and processes, is unproductive as there are relevant differences between internal and external states and processes. Instead, following Sutton (2010), we should conceive of the internal and external components of situated cognitive systems as complementary. When arguing for complementarity between the internal and the external, it is important to be able to

conceptualize how these internal and external components are integrated into agent-artifact systems that perform cognitive tasks. To this end, I propose a multidimensional framework to conceptualize the degree of integration. In this framework, I synthesize and build on the work of Sutton (2006), Sutton et al (2010), Wilson & Clark (2009), Sterelny (2010), and Menary (2010c).

These theorists have provided some of the relevant dimensions, but tend to prioritize certain dimensions while overlooking others. I refine and synthesize some of their dimensions into a coherent and systematic multidimensional framework, add a number of dimensions to the framework, and examine where and how some of these dimensions overlap and interact. This results in a multidimensional framework in which situated cognitive systems can be located. The dimensions in the framework include epistemic action and information flow, speed of information flow, reliability, durability, trust, procedural transparency, informational transparency, individualization, and transformation. Importantly, these dimensions are not meant as necessary and sufficient conditions for cognitive extension and thus do not provide a clear set of conditions to demarcate between cases of embedded and extended cognition. On my view, which is partly based on that developed by Sutton et al (2010), it is more fruitful to think of situated cognitive systems as populating a certain region in this multidimensional space. The higher a system scores on these dimensions, the more tightly coupled the system is and the deeper the artifact and its user are integrated into situated systems that perform cognitive tasks. This framework thus provides a new perspective on the conditions for cognition extension and distribution.

Finally, in order to put the taxonomy and multidimensional framework to work and to demonstrate their value, I focus on the use of cognitive artifacts in scientific practice, particularly molecular biology laboratories. First, I classify cognitive artifacts in molecular biology laboratories on the basis of the taxonomy, and then conceptualize how and how deeply some cognitive artifacts and their users are integrated into systems that perform cognitive tasks related to scientific practice. One of the reasons for focussing on molecular biology laboratories is because I did an undergraduate degree in this field and thus have firsthand experience with the use of cognitive artifacts in this context. This experience will turn out to be helpful when investigating the micro-interactions between agents and artifacts in laboratories.

A second reason is that the use of artifacts in scientific practice has not been a prominent topic in traditional philosophy of science and is therefore a fruitful topic for exploration. Only recently have philosophers of science became interested in instruments and their functional role in experiments and the creation of knowledge. There are many kinds of scientific instruments exhibiting different properties and functions, and some philosophers have developed classifications in order to create systematicity and to better understand what different instruments do. I review a number of current classifications of scientific instruments (Baird 2003, Harré 2003; Nersessian 2005), compare them to my taxonomy, and point out that current classifications overlook a number of important categories of cognitive artifacts and thus cannot account for all cognitive artifacts in molecular biology labs. The taxonomy I develop can account for all cognitive artifacts in molecular biology labs.

A third reason is because there is a small body of literature that is concerned with distributed cognition and scientific practice (Giere 2002a, 2002b; Magnus 2007; Nersessian 2005, 2009). I critically engage with this body of literature and compare it with my framework for conceptualizing complementary integration of agents and artifacts. I argue that although these approaches are perceptive, they focus exclusively on representational artifacts and are thus somewhat limited in scope. They would also benefit from a more detailed look at the micro-interactions between researcher and artifact, as these determine how extended/distributed/integrated a situated cognitive system is. The multidimensional framework is able to investigate these micro-interactions and is thus a valuable addition to current approaches. The last step in my analysis is to conceptualize in which ways cognitive artifacts in molecular biology labs and their users are integrated into extended/distributed cognitive systems by applying the multidimensional framework to four case studies: (1) computer models of protein folding, (2) pH-meters, (3) laboratory notebooks, and (4) organised workplaces.

3. Methodology

The overall goal of this thesis is to better understand the variety of situated cognitive systems consisting of embodied agents and cognitive artifacts, and to conceptualize how such artifacts and their users are integrated into systems that perform cognitive tasks (in scientific practice). As this goal is mostly conceptual in nature, the methodology used to achieve this goal is conceptual analysis. Particularly, the concepts

of artifact, technique, embodiment, function, information, representation, epistemic action, cognitive task, scientific instrument, models, and scientific practice are central and analysed in different chapters throughout this thesis. These concepts are typically discussed by different fields in philosophy. Artifact, technique, and function are discussed by philosophy of technology; embodiment, information, representation, epistemic action, and cognitive task are discussed by philosophy of cognitive science; and scientific instrument, models, and scientific practice by philosophy of science. Various fields in philosophy thus contribute to achieving the overall goal of this thesis.

By using and further developing these concepts, this thesis aims to contribute to distinct debates in these fields, but also aims to bring these debates into contact and to examine how and where these concepts and debates can be synthesized and cross-fertilize. For example, philosophy of technology can contribute to situated cognition theory by providing the relevant concepts to better understand the metaphysical, informational, and functional properties of cognitive artifacts. Better understanding these properties provides a better grip on how artifacts are used and integrated into the cognitive system of their users. Situated cognition theory can contribute to the debate on scientific instruments and model-based reasoning (see also Nersessian 2008) by providing relevant concepts for better understanding how instruments and models relate to and are integrated with the cognitive systems of their users. Philosophy of technology can contribute to the debate on instruments and models in philosophy of science by providing the concepts to better understand their artifactuality and functions (see also Brey 2003). The research in this thesis thus contributes to distinct debates but also builds intradisciplinary bridges between different fields within philosophy.

In addition to conceptual analysis, synthesis, and building intradisciplinary bridges, this thesis also draws on empirical research from cognitive science and archaeology to ground some of the conceptual claims. Throughout this thesis, I draw on work in distributed cognition which is grounded in anthropological and ethnographic research (Hutchins 1983, 1995a, 1995b; Kirsh and Maglio 1994; Kirsh 1995; Nersessian 2005, 2009). The concept of human embodiment and the related notion of body schema, as discussed in chapter 2, is grounded in empirical work in cognitive science (e.g. Johnson-Frey 2003; Maravita & Iriki 2004). The methodology for taxonomizing cognitive artifacts, as developed in chapter 4, draws on artifact classification in archaeology

(Adams & Adams 1991). Finally, in the concluding chapter, I propose a methodology to empirically investigate the degree of integration between agents and artifacts. This methodology complements the conceptual framework developed in chapter 5.

4. Contributions to the Field

The main contributions of this thesis are intended to be:

- An analysis, synthesis, and critique of current characterizations and classifications of cognitive artifacts (chapter 2).
- A general, high-level classification of the components of situated cognitive systems, including those that are artificial (i.e., artifacts and techniques), natural, and social (chapter 2).
- A notion of cognitive artifacts as a functional kind, including those with proper and system functions (chapter 3).
- An information-centered approach to taxonomizing the artifactual elements in situated cognitive systems, based on artifact categorization in archaeology. This results in a rich and detailed taxonomy, outlining and defining a variety of distinct but sometimes overlapping categories of cognitive artifacts, including those exhibiting representational and ecological information (chapters 3 & 4).
- A multidimensional framework for conceptualizing the complementary integration between embodied agents and cognitive artifacts (or other external resources). This framework provides a new perspective on the conditions for cognitive extension or cognitive distribution by arguing that we should not look at cognitive extension/ distribution in terms of necessary and sufficient conditions, but as an inherently multidimensional phenomena (chapter 5).
- An analysis and critique of current classifications of scientific instruments (chapter 6).
- A classification of cognitive artifacts used in molecular biology laboratories in terms of their informational properties (chapter 6).
- An analysis and critique of current approaches to distributed cognition in scientific practice (chapter 6).
- A number of case studies of the complementary integration of embodied agents and cognitive artifacts in molecular biology laboratories, providing a better

understanding of scientific practice concerning the use of cognitive artifacts (chapter 6).

5. Structure of the Thesis

This thesis has two parts. Part I is mainly concerned with the individual components of situated cognitive systems and starts by identifying and classifying such components, including those that are artificial (i.e., artifacts and techniques), natural, and social (chapter 2). It then narrows in on cognitive artifacts by developing a detailed taxonomy, outlining a number of distinct but sometimes overlapping categories of cognitive artifacts which are based on the informational properties of the artifact. This taxonomy is developed by drawing on artifact categorization in archaeology (chapters 3 & 4).

Part II is mainly concerned with situated cognitive systems and starts by presenting a multidimensional framework to conceptualize how artifacts and embodied agents are integrated into situated cognitive systems (chapter 5). It then applies the taxonomy and multidimensional framework to first classify cognitive artifacts in molecular biology laboratories and second to conceptualize how some of these cognitive artifacts and their users are integrated into systems that perform cognitive tasks (chapter 6). This thesis ends by presenting an overall conclusion and pointing out some possibilities for further conceptual and empirical research (chapter 7).

Part I

Components of Situated Cognitive Systems

2

Classifying the Components of Situated Cognitive Systems

1. Introduction

The goal of this chapter is to develop a preliminary classification of the components of situated cognitive systems[1]. Some current classifications (Donald 1991; Norman 1993; Brey 2005; Sterelny 2004; Nersessian et al 2003) focus exclusively on representational artifacts. However, given that embodied agents also use ecological artifacts, cognitive techniques, natural objects, and other people to help them perform their cognitive tasks, these classifications give an incomplete picture of the diversity of the components of situated cognitive systems. Moreover, not only do they give an incomplete picture, their methodology for classifying representational artifacts is cognition-centered in that it takes a human agent and its cognitive processes and goals as point of departure and then classifies artifacts on the basis of the cognitive process or goal to which the artifacts contribute. Such cognition-centered approaches do not pay a great deal of attention to the informational properties of the artifactual element in situated cognitive systems, because that is not what they are trying to explain. In order to develop a more inclusive classification, this chapter presents a preliminary

[1] A brief note on terminology is helpful here. I use the words "classification" and "categorization" interchangeably to indicate a method or system of distinguishing between different entities and to group those entities in categories based on some shared property. A taxonomy is a particular classification/categorization system that has categories (or taxa) on more than one level of abstraction. Taxonomies are thus hierarchical and multi-level classification/categorization systems. More on this in section 3.2 of chapter 3.

taxonomy of the components of situated cognitive systems and points out some limitations of current approaches.

The chapter is organized as follows. It starts by reviewing and discussing a number of suggested characterizations and classifications of cognitive artifacts. In discussing these views, I distinguish between artifacts, naturefacts, techniques, and social scaffolds (section 2). Next, I further conceptualize different categories of artifacts and techniques. I also contrast embodied artifacts with cognitive artifacts, arguing that these are not mutually exclusive categories (section 3). Lastly, a concluding summary is given (section 4).

2. A Review of Suggested Characterizations and Classifications

This section surveys the literature on the components of situated cognitive systems. The emphasis in the current literature is on artifacts and this emphasis is reflected in this section.

2.1 Edwin Hutchins

In his book, *Cognition in the Wild*, cognitive anthropologist Hutchins (1995) is at pains to avoid developing a category of cognitive artifacts[2]. He writes that:

"We are all cognitive bricoleurs – opportunistic assemblers of functional systems composed of internal and external structures. In developing this argument I have been careful not to develop a class, such as cognitive artifacts, of designed external tools for thinking. The problem with that view is that it makes it difficult to see the role of internal artifacts, and difficult to see the power of the sort of situated seeing that is present in the Micronesian navigator's images of the stars" (Hutchins 1995, p. 172).

Hutchins' worries here are twofold: if we focus on cognitive artifacts as the most important component of situated cognitive systems, it makes it difficult to see the functional roles of (a) internal artifacts and structures, and (b) of external structures that are not human-made. To make his case, Hutchins points out that Micronesian navigators use the stars as material anchors to navigate at sea (see also Hutchins 2005).

[2] In fact, Hutchins only mentions the term "cognitive artifact" once in his entire book, which is in the above quote.

He argues that the interaction between internal artifacts, or perceptual strategies, and the external stars makes it possible for the Micronesians to navigate. Thus, in order to explain the navigational capacities of the Micronesians, we have to take into account their learned perceptual strategies and the stars, neither of which are proper artifacts. If we only focus on situated systems consisting of embodied agents and cognitive artifacts, we might overlook interesting cases like these, which would reduce the scope of situated cognition theory. And, furthermore, if we only focus on the artifactual component in such situated systems, we might overlook the functional role of what Hutchins calls "internal artifacts".

I am sympathetic to Hutchins' worries, but they can be overcome by acknowledging and emphasizing that we should be aware that artifacts are only one possible component of particular situated cognitive systems that interact and are integrated with other components[3]. Developing a category of cognitive artifacts does, in my view, not mean that other components of situated systems are ignored or overlooked. Ultimately, we should study situated cognitive systems, rather than their components, but this does not mean that we cannot develop categories and vocabularies for their components, as this would equally reduce scope.

Before I continue outlining Hutchins' account, I address two terminological issues by distinguishing between different components of situated systems. First, it is clarifying to make a distinction between technology and technique. A piece of technology (or artifact) is typically defined as a physical object or structure intentionally designed, made, and used for a particular purpose[4], whereas a technique (or skill) is a method or procedure for doing something. Techniques may involve interacting with artifacts. A car, for instance, is an artifact and the way I interact with it to drive is a technique.

---

[3] How cognitive artifacts and their users are integrated into systems that perform cognitive tasks is conceptualized in chapters 5 and 6.

[4] This definition is sufficiently broad as to include less prototypical cases of artifacts such as domesticated animals (e.g., guide dogs) and genetically modified organisms (e.g., biofuel producing algae). For the latter, see Sune Holm (2013), and see Dan Sperber (2007) for a general discussion on these issues. Guide dogs and biofuel producing algae are intentionally modified (or trained) by humans to perform a particular function, i.e., guiding blind people or producing biofuel. Thus the material of which (cognitive) artifacts are made can be biological or non-biological and in some cases (cognitive) artifacts may even be alive, e.g., in the case of a guide dog.

Likewise, a photo camera is an artifact and the way I interact with it to take photographs is a technique. So quite often (though certainly not always) techniques concern interacting with artifacts. Both technologies and techniques are intentionally designed and used for some purpose and are in that sense artificial, i.e., human-made. Technologies and techniques are also both "for something" and are thus functional entities. However, it is important to note, or so I claim, that they are not both artifactual. Only technologies are artifactual in that they are designed and manufactured physical objects or structures and in this sense what Hutchins refers to as internal artifacts, such as perceptual strategies, can best be seen as cognitive techniques, rather than as internal artifacts. I think it is better to reserve the term "internal artifact" for designed, physical objects that are implanted in the human body such as pace makers, retinal implants, cochlear implants, invasive brain-computer interfaces, and other internal artifacts. Moreover, given that these navigation techniques are learned from other navigators and are thus first external to the organism, it is perhaps more accurate to refer to them as internalized, rather than as internal.

Second, Hutchins writes that Micronesian navigators use the stars in the same way as manufactured navigational artifacts are used, loosely implying that they are a kind of cognitive artifact. Whilst stars are neither artificial nor artifactual in the sense just explained, they are nevertheless the perceptual object of a cognitive technique and have a functional role in navigation. Using natural objects or structures for some purpose is not uncommon, for example, using a dead branch of a tree as a walking stick, a stone as a hammer, or, indeed, the stars to navigate. When doing so, the branch, stone and stars are not intentionally made for those purposes and may be seen to form a bridge between natural objects and artifacts. Risto Hilpinen (2011), in his entry on artifacts in the *Stanford Encyclopedia of Philosophy*, refers to such objects as naturefacts. So, following Hilpinen's terminology, I suggest referring to natural objects that are used for cognitive purposes as cognitive naturefacts. This does not make them less important for performing cognitive tasks. I am not privileging cognitive artifacts in any way. I am just saying that because stars are not intentionally designed, made, or modified for some purpose, they do not belong in a category of artifacts. Hutchins' example is apt in that it shows that humans as cognitive agents not only intentionally construct and modify their cognitive niche, but even exploit natural objects for cognitive purposes.

Having addressed these terminological issues and distinguished between cognitive artifacts, cognitive techniques, and cognitive naturefacts, let me continue outlining Hutchins' account. Although he developed neither a category nor a taxonomy of cognitive artifacts, his cognitive ethnographic study of ship navigation does contain a section, *Sources of Information for Position Fixing*, in which he describes external navigational artifacts and their representational and functional properties in quite some detail. A Hutchins-style analysis of cognitive artifacts takes an ethnographic approach to identify and study, but not to categorize, how artifacts play distinct functional roles in cognitive tasks. There is deliberately no attempt to systematize or categorize these artifacts, because of the above mentioned reasons.

In *Cognition in the Wild,* Hutchins was reluctant to develop a category of cognitive artifacts. However, in his entry in the *MIT Encyclopedia of the Cognitive Sciences* he is less reluctant to give a definition. He writes that "cognitive artifacts are physical objects made by humans for the purpose of aiding, enhancing, or improving cognition" (Hutchins 1999, p. 126). The conditions in this definition specify that cognitive artifacts are human-made, physical objects that positively impact human cognition. He continues by giving examples. Drawing on King Beach (1988), he mentions that bartenders sometimes remember the sequence of the ordered drinks by structuring distinctively shaped drink glasses such that they correspond to the sequence of the ordered drinks. As a result, the bartenders do not have to remember the order of the drinks, but just offload it onto their environment. Beach refers to the structuring of the drink glasses to offload memory as "material mnemonic symbols".

A second example Hutchins mentions, drawing on David Kirsh (1995), is the systematic arrangement of ingredients and cooking equipment while preparing a meal. The ingredients and equipment are arranged such that they facilitate the cooking process by offloading the order of the steps of the cooking process onto the environment. Kirsh refers to this as "intelligent use of space" in which the arrangement of artifacts is itself a cognitive artifact. Other examples Hutchins mentions are a string tied around one's finder as a reminder, calendar, shopping list, calculator, and computer. These are all human-made, physical objects that improve cognition, including memory, decision-making, planning, numerical cognition, and other cognitive processes. The above mentioned examples show that "there is a continuum from the case in which a cognitive artifact is used as designed, to cases of cognitive uses of artifacts that were made for

other purposes, to completely opportunistic uses of natural structure" (Hutchins 1999, p. 127)[5].

Hutchins concludes his encyclopaedia article with the following paragraph:

"There is no widespread consensus on how to bound the category of "cognitive artifacts." The prototypical cases seem clear, but the category is surrounded by gray areas consisting of mental and social artifacts, physical patterns that are not objects, and opportunistic practices. The cognitive artifact concept points not so much to a category of objects, as to a category of processes that produce cognitive effects by bringing functional skills into coordination with various kinds of structure" (Hutchins 1999, p. 127).

Hutchins' conclusion is that we should not be looking at external artifacts alone, nor should we be looking at cognitive techniques (or other internal processes) alone, but at interactive processes between functional skills and different kinds of physical structure (including the prototypical cases of cognitive artifacts)[6] that produce cognitive effects. In other words, we should be studying situated cognitive systems and not their components. I fully agree with Hutchins that we should not neglect the importance of cognitive techniques and that the interaction between functional skills and material structure is essential for the study of cognition. However, I do not see why developing distinct categories for different processes and objects (including cognitive artifacts) that produce cognitive effects would discourage us from looking at the interactivity and integration between internal and external processes and structures. There is no a priori reason why cognitive artifacts, cognitive naturefacts, and cognitive techniques cannot be subcategories of a more general category of cognitive scaffolds.

---

[5] See also Preston (2013) for an excellent analysis of how artifacts can be designed or improvised, resulting in different kinds of artifact functions. These different kinds of artifact functions are outlined in the next chapter.

[6] The prototypical cases that Hutchins refers to seem to be artifacts with straightforward representational functions such as calendars, shopping-lists, maps, and radar systems. Less prototypical cases are what Beach refers to as "material mnemonic symbols" and what Kirsh refers to as "the intelligent use of space". These are artifacts that have initially not been designed to aid cognition, but have cognitive functions because a human agent has attributed a cognitive function to such artifacts during improvisation. The design and improvisation of cognitive functions is explained in detail in the next chapter.

Contrary to Hutchins' view on the category of cognitive artifacts, I think there are at least three reasons why it is important to develop a special subcategory of cognitive artifacts and to distinguish between and taxonomize different kinds, based on their informational properties. First, as Hutchins rightly pointed out, we are opportunistic assemblers of functional systems that are composed of internal (or internalized) structures and external structures. From an agent-centered perspective, it does not matter whether these external structures are artifactual or natural. What matters is that they functionally contribute to performing a cognitive task. So, in one sense, artifacts and naturefacts are continuous in that they can both function as external cognition-aiding structures. There is, however, one relevant difference between artifacts and naturefacts that justifies paying more attention to (taxonomizing) artifacts: our intentional control over the informational content and functions of cognitive artifacts is considerably larger and results in significantly more variety, as compared to the intentional control over cognitive naturefacts and the resulting variety. We have intentional control over the content and functions of cognitive naturefacts only insofar as we can choose which natural objects or structures to use for some cognitive task.

One could argue that we also do not have full intentional control over the content and functions of all cognitive artifacts, as some are designed and made by others. Maps, timetables, textbooks, manuals, encyclopaedias, and roads signs, for example, are designed and made by agents outside the situated system and a user typically has no control over the content of such artifacts. The informational content of other cognitive artifacts, however, *is* designed and made by the user of the artifact. Notebook and diary entries, to-do lists, shopping-lists, PowerPoint slides, and an architect's sketch, for example, are typically made by the user of the artifact. So the intentional control an agent has over the content and functions of a cognitive artifact differs, depending on the kind of artifact (Heersmink 2013b, see also chapter 4 of this thesis). But, in either case, cognitive artifacts are intentionally designed and made to aid human cognition. The intentional design and making of cognition-aiding artifacts (either by designers or users) results in a broad range of cognitive artifacts exhibiting different kinds of informational properties that are specifically geared towards realizing a broad range of cognitive tasks, including navigating, calculating, remembering, measuring, planning, designing, etc. This kind of intentional control not only results in a much richer variety

of cognitive artifacts, as compared to cognitive naturefacts, but also results in external artifactual structures that can be integrated much deeper into the onboard cognitive system, because they are functionally and informationally malleable. Consequently, the transformative impact of artifacts on our cognitive system and practices, both ontogenetically and phylogenetically, seems much more substantial as compared to naturefacts.

Second, it gives us a much deeper conceptual understanding of a particular kind of artifact, namely, cognitive artifacts. This is important because it contributes to expanding and further developing a relatively small and emerging subfield in the philosophy of technology, which is sometimes referred to as philosophy of artifacts. As Randall Dipert (1993) has argued, an adequate philosophical theory of artifacts is largely lacking in the history of analytical Western thought (compare Houkes & Vermaas 2010; Kroes 2012; Preston 2013). Given the ubiquitousness of artifacts and their fundamental role in our lifeworld, culture, and cognition (Preston 1998a), an adequate philosophical theory of artifacts would be very much welcome. Developing a taxonomy of cognitive artifacts contributes to such a theory, while also being beneficial for situated cognition theory.

Third, the reason cognitive artifacts are important to better understand is because they have different (i.e. complementary) properties and affordances as compared to cognitive techniques and other internal cognitive states and processes. Although such states and processes are also physical in that they supervene on neural and sometimes bodily structures, the physical material of artifacts is very different and allows operations that are very difficult to perform in the brain. The specific physicality and operations external artifacts allow, gives them particularly distinct functional and informational properties, which are important to study in their own right. As Donald (1991) has pointed out, exograms (or external representational systems) have properties that are different from engrams (internalized information in biomemory). Engrams are internalized and realized in the medium and format of the brain, whereas exograms are external and much less constrained in their format and capacity. The storage capacity of exograms far exceeds the storage capacity of both single entries and clusters of entries in biological memory. Exograms are flexible in that they can be reformatted and easily transmitted across different media, whereas engrams are less

flexible. These differences certainly do not always apply, but when they do apply, they are enabled by the particular physicality, malleability, and format of external artifacts.

It is precisely because exograms have such different functional and representational properties, as compared to engrams, they have the capacity to *complement* the properties of engrams (Sutton 2006, 2010). This is arguably why human beings have developed exograms in the first place and if we want to understand how they complement the working of engrams, cognitive techniques, and other cognitive states and processes, it is very helpful to have a taxonomy that distinguishes between different types of cognitive artifacts, outlining their distinctive cognition-aiding properties.

For these reasons, an important first step towards a better understanding of some situated cognitive systems, I claim, is a taxonomy of external cognitive artifacts. The very reason we design, produce, and deploy external cognitive artifacts is because internal states and processes have limited capacity and are limited in scope. In order to achieve our cognitive goals, we regularly deploy external artifacts, thereby complementing the limitations of our onboard cognitive system. Successful epistemic interaction requires fine-tuning and integration between technique and artifact. A better understanding of the artifactual element in a situated cognitive system can give a better understanding of the epistemic, interactive, and integrative process between technique, artifact, and agent. Focussing on the informational properties of artifacts allows us to see which cognitive capacities and functional skills an agent ought to have to successfully deploy the artifact. So, it allows us to sketch the contours of one piece of the puzzle, which in turn, allows us the better understand the contours of the other pieces in the puzzle, i.e., embodied brains and their cognitive techniques.

2.2 Exograms as Symbolic Technologies

I start by discussing Merlin Donald's notion of exograms. During the evolution of our cognitive system, we have developed a variety of notational systems which allow us to offload and store information in our environment, thereby creating an external memory field. Donald writes that: "The external memory field usually consists of a temporary array of visual symbols immediately available to the user. The symbols are durable and may be arranged and modified in various ways, to enable reflection and further visual processing" (Donald 1991, p. 296-297). Those external symbols are referred to by

Donald as exograms, as opposed to engrams. "An exogram is simply an external memory record of an idea", whereas an engram is "a single entry in the biological memory system" (Donald 1991, p. 314).

Donald explains that exograms have different properties as compared to engrams. Engrams are internalized and realized in the medium and format of the brain (which is typically visual/experiential or linguistic/propositional), whereas exograms are external to the brain and much less constrained in their format. The storage capacity of exograms far exceeds the storage capacity of both single entries and clusters of entries in biological memory (compare, for example, an encyclopaedia with the knowledge an individual has about the meaning of words and concepts). Exograms are much more flexible in that they can be reformatted and easily transmitted across different media, whereas engrams are less flexible. The most important property of exograms, Donald argues, is their capacity for continuous refinement. Exograms are human-made and have undergone and are undergoing a process of iteration, testing, and improvement. They allow us to externalize and freeze the products of thinking in time and to examine and change their content in an iterative process, which is very difficult to do in the brain (Donald 1991, p. 314-316).

Donald sketches some useful dissimilarities between the properties of exograms and engrams, but perhaps some of them are a bit too extreme (see also Sutton 2010, p. 206-207). These dissimilarities suggest that engrams are always fixed, constrained, limited, static, and unreliable, whereas exograms are flexible, unconstrained, reformattable, unlimited, highly dynamic, and so on. It is important to note that there are certainly exceptions to this general contrast that Donald sketches, as some particular exograms are in fact fixed, contain a small amount of information, are limited, and unreliable, whereas certain engrams are flexible, contain a large amount of information, are dynamic, and reliable. For example, a token map of London may be more static, less up-to-date, and may contain less information, as compared to a taxi driver's engram of the layout of London's streets. London taxi drivers have developed an unusual capacity to memorize the complex layout of London's streets (Maguire et al. 2000) and so their engrams of London's layout, i.e., their cognitive maps, are larger in the amount of information they contain than most available tourist maps, are constantly updated and thus more accurate and flexible than most maps, and less constrained than most maps.

So there are certainly cases where engrams are significantly more flexible and dynamic as compared to exograms.

More recently, Donald has referred to exograms as "symbolic technologies", which are "specifically designed to represent, communicate and store knowledge. Such objects introduce a completely new element into human cognition: external, that is, non-biological, memory storage (as in an encyclopedia, for example)" (Donald 2010, p. 71). Symbolic technologies thus have communicative, representational, and knowledge storing functions, thereby (in a sense) relieving and transforming the human memory system. Donald's main point is, therefore, that a human cognitive system including exograms has radically different (memory) properties than a merely biological one. In short, exograms are physical, external memory devices that have transformed the nature of our memory as well as our cognition more generally.

Significant objects, amulets, totems, masks, magical tokens
Transient and permanent iconography (in sand, mud, stone)
Crafted memory devices such as knotted cords
The built environment
Painted and sculpted images, such as cave paintings and totems
Astronomical measuring devices, e.g. stone circles and burial mounds
Trading tokens
Early scripts for trade and crop administration
Longer written records of crops, laws, edicts, genealogies
Works of literature, poetry
Mathematical and geometrical notations
Architectural and engineering drawings, models
Libraries and archives
Elaborate scientific and navigational instruments
Moving pictures, computers, electronic media
'Smart' machines, robots, high-tech virtual environments

Table 1. Generic exogram systems in terms of their historical emergence.

Table 1 (taken from Donald 2010, p. 72) contains an impressive list of generic exogram systems: from amulets to poetry, from stone circles to written laws, and from cave

paintings to virtual environments. It is, however, a mere list of examples of different exogram systems and the only structure and systematicity is that it is roughly chronological in terms of emergence, with some historical overlap. So Donald here loosely categorizes exogram systems on the basis of their *historical properties,* i.e., on the basis of the historical order in which they have emerged. Note that there is no categorization based on the functional or informational properties and capacities of exograms. One of Donald's goals is to conceptualize the historical-evolutionary influence of exograms on our cognitive system. Given this goal, an historical classification is of course important. However, one of the goals in Part I of this thesis is to develop a systematic taxonomy of external cognitive artifacts (including exograms) that are currently used. On the basis of their functional and informational properties, I develop a set of categories in which cognitive artifacts with similar properties can be grouped. So rather than giving an historical classification, I aim to develop a framework in which examples can be grouped into categories defined by their functional and informational properties, not their historical properties.

2.3 Cognitive Artifacts

The phrase "cognitive artifact" has seen a number of characterizations in the situated cognition literature and I discuss four of them below.

2.3.1 Donald Norman

An early and quite influential definition of cognitive artifacts has been put forward by psychologist and design theorist Donald Norman. He argued that cognitive artifacts are "artificial devices that maintain, display, or operate upon information in order to serve a representational function and that affect human cognitive performance" (Norman 1991, p. 17). So, according to this definition, a cognitive artifact is an artificial device that is designed to represent information that has an impact on human cognition. Norman is thus particularly interested in artifacts containing external representations, or "surface representations" in his terminology, including written language in books and on blackboards, icons on a desktop, diagrams, Roman and Arabic numerals, graphs, checklists for pilots, abacuses, and even indentions in clay, sand, or wood. He argues that, in order to understand the interactions between and complementary characteristics of humans and cognitive artifacts, we should study the properties of both cognitive artifacts as well as the human agents that use them. In his words:

"The study of the artifact informs us about the characteristics of the human. The study of the human informs us of the appropriate characteristics of artifacts. And the study of both the artifact and the human must emphasise the interactions between and the complementarity of the two" (Norman 1991, p. 16).

In his book, *Things that Make us Smart: Defending Human Attributes in the Age of the Machine,* he expands his notion of cognitive artifacts, to include mental artifacts. He writes:

"We humans have invented a wide variety of things to aid our cognition, some physical, some mental. Thus tools such as paper, pencils, calculators, and computers are physical artifacts that aid cognition. Reading, arithmetic, logic, and language are mental artifacts, for their power lies in the rules and structures that they propose, in information structures rather than physical properties. Mental artifacts also include procedures and routines, such as mnemonics for remembering or methods for performing tasks. But whether physical or mental, both types of artifact are equally artificial: They would not exist without human invention. Indeed, anything invented by humans for the purpose of improving thought or action counts as an artifact, whether it has physical presence and is constructed or manufactured, or whether it is mental and taught" (Norman 1993, p. 4-5).

According to Norman, a substantial part of our problem-solving skills results from our capacity to create artifacts, both mental and external. But although he acknowledges the existence of mental artifacts whose power lies in information-structures rather than physical properties, his explanatory project focusses on physical, external artifacts. In his words: "Without external aids memory, thought, and reasoning are all constrained. The real powers come from devising external aids that enhance cognitive abilities" (Norman 1993, p. 43). Throughout his book, he emphasizes that external cognitive artifacts complement the cognitive capacities of an unaided human mind. They do so by representing information in a way that the brain is not capable of or finds difficult to do. Therefore, he concludes, "the power of cognitive artifacts derives from the power of representation" (Norman 1993, p. 75). Thus, like Donald, Norman thinks that what really matters are representational cognitive artifacts, including written language, icons, diagrams, and numerals. Moreover, an essential property of the representations

supported by cognitive artifacts, Norman points out, is that those representations are themselves also artificial objects that can be perceived and studied.

Norman groups cognitive artifacts into two categories: experiential and reflective artifacts. Before discussing these two types of artifacts, it is helpful to say a few words on the different modes of thought they aid, which are experiential and reflective cognition. Norman presents these modes of thought as follows. Experiential cognition concerns efficiently and effortlessly perceiving and reacting to our environment and is the cognitive foundation for expert skills. It is often concerned with routine tasks, including those of expert pilots, mechanics, or athletes. Making decisions in this mode does not require reflective planning or explicit problem solving, but is rapid, transparent, somewhat reflex-like, reactive, and comes without much conscious effort. "The experiential mode of performance is one of perceptual processing: what cognitive science calls pattern-driven or event-driven activity" (Norman 1993, p. 26). So our perceptual systems take patterned information onboard, process it, and then allow us to rapidly and fluently respond to events in the environment.

The reflective mode, by contrast, is more concerned with deliberate reasoning, conscious decision-making, making inferences, and comparing and synthesizing (new) ideas. Reflective cognition is often slow, laborious, and requires the ability to store intermediate results, to make inferences from stored information, evaluate premises, and construct arguments. "In terms of cognitive science, reflective cognition is conceptually driven, top-down processing" (Norman 1993, p. 25). So it is more concerned with conceptual thought than with skilfully reacting to patterns or events in the environment. Norman warns us that these two modes of thought are not meant to capture all of human cognition, which is a much more multidimensional activity, they are also not independent and quite often overlap and are thus not mutually exclusive. However, from a practical point of view, the distinction between experiential and reflective cognition is worth considering, he says, "in part because much of our technology seems to force us towards one extreme or the other" (Norman 1993, p. 26). Norman's distinction between experiential and reflective cognition is reflected in current dual-process theories in cognitive science (e.g. Evans 2008; Evans & Stanovich 2013). In such approaches to cognition, there is a need to understand the exact relation between the two modes of thinking. Perhaps the way that Norman links his dichotomy with an account of cognitive artifacts might enrich dual-process theories which are

often highly individualistic in that they focus only on cognitive processes occurring in the brain.

Experiential and reflective cognition are supported by different types of artifacts, exhibiting different representational properties.

"Tools for experiential cognition should make available a wide range of sensory stimulation, with enough information provided to minimize the need for logical deduction. Similarly, tools for reflection must support the exploration of ideas. They must make it easy to compare and evaluate, to explore alternatives. They should not restrict behavior to the experiential mode" (Norman 1993, p. 26).

"Experiential artifacts have different functions from reflective ones. Experiential artifacts provide ways to experience and act upon the world, whereas reflective artifacts provide ways to modify and act upon representations" (Norman 1993, p. 52).

These two types of cognitive artifacts are thus distinguished by their function. The function of experiential artifacts is to allow us to indirectly experience events, things, or phenomena in the physical world and provide us with information that is usually inaccessible to our perceptual systems. They provide us with a semi-transparent informational window on some aspect(s) of the physical world. Examples Norman mentions are: a telescope, which gives us information about things in (far) space; a recording, which allows us to experience something that happened in the past or in some other part of the world; and measuring instruments such as the gas gauge, which give us information about states of equipment. Such artifacts perceptually mediate between the human mind and some aspect(s) of the physical world and the information they provide is ideally rather straightforward and easy to interpret, as to minimize the need for logical, reflective thought.

By contrast, reflective artifacts allow us to disregard the physical world and to focus on the representation itself. Telescopes, recordings, and measuring instruments, directly represent the world as it is or was, but reflective artifacts contain representations that can be modified, as to facilitate a problem-solving task. We do our thinking while we modify (parts of) the representation and the physical world it represents is temporarily bracketed off. Thinking with the aid of such dynamic representations, Norman argues,

allows us to discover higher-order relationships, structures, or (in)consistencies in the world. So their function is not so much to directly represent the physical world, but to facilitate and aid reflective, higher-order cognition by providing ways to modify and act upon representations.

One example of a reflective artifact that Norman put forward is the use of everyday objects such as pencils and paperclips as *stand-ins,* as to explain how a car accident happened. One pencil stands in for a car that was hit in the back, a second pencil stands in for a car that hit the first car in the back, a third pencil stands in for a car that hit both other cars from the side, and a paperclip stands in for a dog that ran across the street, causing the first car to hit the second. Using physical objects to represent a certain event creates a representational structure that aids our capacity to describe certain events, thereby enabling others to better understand the situation and making it easier to analyse possible alternative actions. For the purpose of explaining the car accident, it does not matter that the objects do not resemble cars or dogs in any obvious way, the only thing that matters for explaining *this particular event* is their location in relation to the other objects (Norman 1993, p. 47-49).

Norman presents a *cognition-centered* approach to categorizing cognitive artifacts. He takes experiential and reflective cognition as a point of departure and then groups artifacts into two categories, in one category he puts artifacts that aid experiential cognition and in the other category artifacts that aid reflective cognition. Category membership is thus solely based on the mode of cognition the artifact aids, not on representational properties of the artifact. In other words, an artifact is categorized depending on which cognitive systems process, deploy, and are aided by the information that the artifact represents. Consequently, if experiential cognitive systems are aided by the information, it is an experiential artifact, and if reflective cognitive systems are aided by the information, it is a reflective artifact. However, because some artifacts provide information to both systems and can therefore aid both modes of cognition, the same cognitive artifact may, depending on its use, be categorized in both categories.

I illustrate this by developing the example of a team of experienced pilots may need to descend and prepare for landing. In order to decide when to start the descend, the captain and co-pilot will look at the radar screen to see at which longitude, latitude, and

altitude they are flying and to see if there are any possible obstacles in between their current location and the airport. Radar systems are cognitive artifacts that typically aid experiential cognition, as they provide pilots with information that is inaccessible to their perceptual systems and is needed to react to patterns or events in the environment, thereby making adequate and rapid decisions during routine tasks. But the same radar system may also aid reflective thought, for example when a pilot needs to alters its course and quickly plan an alternative route, taking into account their current longitude, latitude, and altitude, their final destination, as well as possible obstacles such as other airplanes, military zones, mountains, and dangerous weather conditions that are represented on the radar screen. Planning this alternative route on the basis of the information the radar represents requires deliberate reasoning, making explicit inferences, and conscious decision-making, i.e., it provides information to reflective systems and thus requires and aids reflective thought. Although the representational properties of the radar system are in both cases identical, the functional role of those properties differ in each mode of cognition. A radar system may thus aid both experiential and reflective cognition.

Similarly, other cognitive artifacts such as diaries also have representational properties that can aid both experiential and reflective cognition. Whilst diaries do not provide a semi-transparent informational window on the physical world in the same way as telescopes, recordings, and measuring instruments do, they can still aid experiential cognition (as well as reflective cognition). It aids experiential cognition when I quickly look in my diary to see whether I am available for a meeting (quite a routine task for most people) and it aids reflective cognition when I use it to carefully plan my teaching timetable, taking into account other obligations, conference visits, and public holidays (not so much a routine task). In the latter scenario, I need to compare and evaluate numerous options, make inferences and conscious decisions in order to construct the timetable. So it requires and aids reflective cognition. Therefore, a token cognitive artifact may have representational properties that aid both experiential and reflective cognition and may thus, depending on its use, be categorized in both categories.

Although Norman's project is greatly concerned with representational properties of artifacts, his method of categorizing cognitive artifacts does not elucidate a lot about the particular representational properties of artifacts in each category, other than that experiential artifacts *usually* provide a semi-transparent informational window on some

aspect(s) of the physical world, whereas reflective artifacts *usually* facilitate interacting with or manipulating of information within an external representation. Analytically, these two conditions are not very helpful because, in some cases, reflective artifacts also provide a semi-transparent informational window on some aspect(s) of the physical world, as illustrated by the above mentioned radar system. So a radar system aids reflective thought not by allowing an agent to interact with or manipulate information, but by providing a semi-transparent informational window on some aspect(s) of the physical world. An artifact may therefore aid reflective cognition without facilitating interaction with or manipulation information.

Conversely, cognitive artifacts that facilitate interaction with or manipulation of information, in some cases, also allow an informational window on the physical world. Norman's example of explaining a car accident with pencils and paperclips is a case in point, as it provides a viewer with a representation of an event. It may not be as semi-transparent as telescope readings, recordings, or measurement outcomes, but it still provides a window on what happened, perhaps in a somewhat similar way as a recording of the same event would have provided. These two conditions, i.e., providing a semi-transparent informational window on some aspect(s) of the physical world and facilitating interaction with or manipulation of information, are therefore not mutually exclusive for category membership. Category boundaries between these two types of cognitive artifacts sometimes overlap.

So, on Norman's view, a cognitive artifact is categorized purely by its functional role in either experiential or reflective cognition, not by representational properties of the artifact, although these may give some indication of category membership. His two-part taxonomy is thus basically a taxonomy of different modes of cognition, not of cognitive artifacts. It is a very helpful heuristic or starting point for exploring how token cognitive artifacts aid experiential and reflective cognition. But if one takes such a cognition-centered approach, it would perhaps benefit from a finer-grained classification, not of *modes* of thought, but of *specific cognitive processes,* including measuring, navigating, memory, calculating, planning, learning, decision-making, and artistic cognition, and to conceptualize how token artifacts aid those specific cognitive processes in particular ways. Such a finer-grained classification is something we will look at in the next subsection.

2.3.2 Philip Brey

In a paper on Human-Computer Interaction, philosopher of technology Philip Brey (2005) borrows and further develops Norman's (1991) definition. He points out that:

"There is a special class of artifacts that are distinguished by their ability to represent, store, retrieve and manipulate information. Norman calls such artifacts cognitive artifacts. He defines them as artificial devices designed to maintain, display, or operate upon information in order to serve a representational function. The keywords here are "information" and "representation." They distinguish cognitive artifacts from other artifacts" (Brey 2005, p. 385).

Thus, like Donald and Norman, Brey argues that cognitive artifacts are artificial devices with representational functions. He uses Norman's definition to develop a brief taxonomy of representational cognitive artifacts on the basis of the type of cognitive process to which the artifact contributes. In his words: "Various classes of cognitive artifacts may be distinguished, based on the primary cognitive capacity or capacities that they extend or aid" (Brey 2005, p. 385). So where Norman made a distinction between experiential and reflective cognition and categorized artifacts as aiding one of those modes, Brey distinguishes between cognitive artifacts that contribute to memory (e.g. notebooks), quantitative interpretation (e.g. thermometers), qualitative interpretation (e.g. colour charts), searching (e.g. labels or search engines), and conceptual thought (e.g. models and diagrams). He has set his sights on Human-Computer Interaction and argues that computers are particularly interesting because they can contribute to all the before mentioned cognitive processes. Computers are also interesting because they can actively and autonomously process information. They are therefore a powerful and versatile cognitive artifact that deserves special attention. I agree with Brey’s conclusion that computers are powerful and versatile cognitive artifacts that deserves special attention, but less so with his approach to classifying cognitive artifacts.

Brey's brief classification is valuable and insightful in that it recognises that there are different types of cognitive artifacts, performing different functions. Like Norman, Brey takes a human organism and its cognitive processes as a point of departure and then conceptualizes which artifacts contribute to which cognitive process. Brey's

cognition-centered approach overlooks and neglects the particular representational properties of the artifact. Whilst Brey's classification is finer-grained than Norman's two-part classification, it has nothing particular to say about the distinctive properties of the artifact, other than that they have representational functions, because (like Norman's classification) it is basically a classification of different types of cognitive processes, not of artifacts. The approach chosen and developed in this thesis (particularly in the next two chapters) takes an *information-centered* approach in that it focuses on informational properties of artifacts and on the basis of those properties it taxonomizes different kinds of cognitive artifacts. So rather than taking modes of cognition or specific cognitive processes as a point of departure for taxonomizing cognitive artifacts, it takes informational properties as a point of departure. One of my aims here is to better understand the range and variety of cognitive artifacts by developing a taxonomy, not of modes of cognition or cognitive processes, but of artifacts. Focussing on the informational properties of such artifacts seems a sound method to develop such a taxonomy. In the next two chapters, this method for taxonomizing is outlined in detail.

2.3.3 John Sutton

In a paper on the different domains and dimensions of distributed cognitive systems, Sutton (2006) develops a concise, high-level taxonomy of resources that contribute to such systems. This taxonomy contains five types of sometimes overlapping resources. First, we deploy cognitive artifacts like sketchpads, notebooks, word-processors, navigational artifacts, and other exogram systems to help us think. Second, we sometimes use natural environmental resources, e.g., "those exploited in ongoing sensory-motor couplings between action and perception" (Sutton 2006, p. 236). Third, Sutton points out that other human agents (or social scaffolds) quite often provide information to achieve our own cognitive purposes. This happens most clearly in transactive memory systems, for example those in dyads (see Wegner 1987; Harris et al. 2011) or larger groups (see Theiner 2013; Williamson & Cox 2013). In such cases of social scaffolding, new forms of memory emerge through the transactive nature of memory. Consider an example of both social and artifactual scaffolding: rally drivers typically rely on a navigator that tells the driver what obstacles to expect, where to turn, the severity of the turn, and so on. This navigational information is usually read by the navigator from previously taken notes. Jointly, driver, navigator, and notebook perform the cognitive task of navigating. Fourth, our embodied capacities and skills can

transform the cognitive task at hand and are sometimes more than just expressions of internally computed thoughts. Examples include embodied interactions with artifacts, gestures, or certain rituals, but also learned skills in sports, music, and dance. As Sutton rightly points out: "These embodied cognitive capacities are interwoven in complex ways with our use of the technological, natural, and social resources mentioned above" (2006, p. 239). And lastly, cases of internal versions of external or cultural resources, i.e., "internalized cognitive artifacts", for example, methods of loci.

Sutton's taxonomy differs from the other taxonomies discussed up to now (Donald, Norman, and Brey), in that it does not only focus on artifacts, but also includes natural environmental resources, other human agents, embodied skills, and internalized cognitive artifacts. Sutton's taxonomy is thus much broader and inclusive and is meant to taxonomize resources or cognitive scaffolds on a higher level of analysis. For this reason, it paints a more realistic picture of the variety of the components of situated cognitive systems.

In the following paragraphs, I focus on what Sutton refers to as "internalized cognitive artifacts". In his words:

"We use a wide range of stratagems to bootstrap, manage, transform, and discipline our minds, and these techniques can coopt internal surrogates as worldly exograms. Both linguistic items - words, labels, phrases - and other symbols can play key cognitive roles independent of any communicative function, in freezing thought or condensing complex affects..." (Sutton 2006, p. 239-240).

Thus internal surrogates such as words, phrases, and other mental representations can be co-opted to function as external exograms usually do, in that way co-opting the properties of worldly exograms into internalized surrogates (see also Sutton 2009b). For example, to memorize the order of the eight planets in our Solar System: Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, and Neptune, we could deploy an external artifact or exogram such as a written list or diagram of the order of the planets. But we could also try to internalize the information an external artifact contains. A common strategy is not to memorize the names of the planets (although that also happens), but to learn a mnemonic phrase such as, for example, "My Very Educated Mother Just Served Us Nachos". The first letter of each word in the phrase corresponds to the first

letter of a planet in the Solar System and (ideally) prompts the recall of the order of the planets. Most people find it much easier to remember one grammatically coherent and meaningful phrase, rather than eight independent and discrete items. Note that the internalized mnemonic phrase has the same functional role in performing a cognitive task as a similar external version, i.e., both have as function to aid the recall of the order of the planets. To give an example: during an open book exam, a student may be asked for (the order of) the planets in our Solar System. In order to answer the question, she may either use a mnemonic that she has written down in her notes or she may use a memorized and internalized mnemonic. Both the external and internalized version of the mnemonic have the same functional role in answering the question. The physical location of the information is irrelevant for answering the question.

Consider a second example of an internalized cognitive artifact, namely, the method of loci (Sutton 2010, see also Hutchins 2005). In short, a cognitive agent memorizes certain spatial relations, for example the rooms in a palace, house, or other building, and has learned to associate certain items, for example images, faces, or lists of words, with particular locations in rooms. When trying to remember the items, the agent imagines walking through the rooms (i.e. the loci) which evokes memories of the associated items. To give an example: when I use the method of loci to construct a shopping list I would image walking through my apartment and put certain items on specific locations: a bottle of wine on my kitchen table, some cheese on my kitchen dresser, milk on the desk in my study room, a baguette near the front door, and so forth. When I am in the supermarket, I would re-imagine the same route through my apartment which would evoke memories of the wine, cheese, milk, and baguette, in that way allowing me to remember what I want to buy. Such methods date back to the ancient Romans and Greek and are still widely used today (see, e.g., Foer 2011). It may seem as if this is quite a burdensome method for memorizing, but, as Sutton points out:

"Despite the apparent doubling of effort required to remember both the locations and then the specific items to be remembered, the system was both economical and flexible, for once the virtual architecture was securely internalized, it could be used and reused at will" (Sutton 2010, p. 62).

So the trick is to firmly internalize spatial relations between rooms and within rooms, what Sutton calls a virtual architecture, which can then be used and reused whenever needed. In both cases of internalized cognitive artifacts (i.e. internalized mnemonics and methods of loci), external information (e.g., linguistic, iconic, or spatial relations between elements) is absorbed by an embodied brain. It is not the case that the cognitive artifact or artifactual environment are literally internalized, only the information they contain is internalized and absorbed. So rather than referring to such phenomena as internalized cognitive artifacts, I suggest referring to them as internalized information. Once that information is internalized, they can be deployed for performing a variety of cognitive tasks (mainly tasks related to remembering) and can be seen as particular cognitive techniques. The relation between internalized information and cognitive techniques is further conceptualized in section 3.5.

2.3.4 Nancy Nersessian

Philosopher of (cognitive) science Nersessian writes: "cognitive artifacts are material media possessing the cognitive properties of generating, manipulating, or propagating representations" (Nersessian 2005, p. 41). She has set her sights on the functional role of cognitive artifacts in scientific and engineering practice and so her examples include scientific models and simulations of phenomena in the world as well as measuring devices that generate representations in visual, quantitative, or graphical format. The particular situated cognitive systems in which she is interested comprise researchers and the cognitive artifacts in a biomedical engineering (BME) laboratory. One example she mentions is a device called the "flow loop", which represents the blood flow in an artery. During a simulation, the flow loop manipulates particular constructs which are representations of blood vessel walls. After the simulation, those constructs are examined with instruments such as the confocal microscope, which generates high-definition images of the constructs. This simulation process allows the researchers to gain information about, for example, the number of endothelial cells and about the direction of filaments in relation to the blood flow. Based on this example, Nersessian concludes, "the *representations generated* by the flow loop *manipulations* of the constructs are *propagated* within the cognitive system" (Nersessian 2005, p. 43, original italics).

In a cognitive-ethnographic study of a BME laboratory, Nersessian et al. (2003, see also Kurz-Milcke, Nersessian & Newstetter 2004; Nersessian 2006; Harmon & Nersessian

2008) present a brief taxonomy of artifacts used in the laboratory. This taxonomy contains three categories: devices (e.g., a flow loop, bioreactor, or bi-axial strain), instruments (e.g., a confocal microscope, flow cytometer, or coulter counter), and equipment (e.g., a pipette, flask, refrigerator, or water bath). Those three categories and the grouping of artifacts in those categories are developed by the BME researchers themselves. This was done during a research meeting with the members of the BME laboratory. On the basis of ethnographic observations, Nersessian and her colleagues have formulated working definitions of the categories developed by the researchers. Devices are defined as engineered fascimiles that serve as in-vitro models and sites of simulation. Instruments generate measured output in visual, quantitative, or graphical format. And equipment assists with manual or mental labour. The authors write that not all artifacts in the taxonomy are cognitive artifacts. "The cognitive artifacts in the distributed systems in the lab cut across these distinctions, though most are devices or instruments. Analysis of the ethnographic data has focussed our attention on the devices, all of which we classify as cognitive artifacts" (Nersessian, et al. 2003, p. 4)

By asking the users of the artifacts to develop categories and to group artifacts in those categories, Nersessian and her colleagues have developed a taxonomy that is interesting and distinctive, in that it uses an empirical, user-centered, and bottom-up approach to categorizing (cognitive) artifacts. It is different from Norman and Brey's cognition-centered approach in which a theorist groups cognitive artifacts in a conceptual, top-down manner. The user-centered taxonomy Nersessian and colleagues present focusses only on scientific practices in one particular BME laboratory and is perhaps hard to apply to all cognitive artifacts. Because one of my aims in this thesis is to develop a picture that can potentially account for all cognitive artifacts, an empirical approach would perhaps be less feasible, as the amount of cognitive artifacts and the contexts in which they are used are quite substantial and therefore difficult to cover during the duration of a PhD project. So, with this particular aim in mind and practical restrictions, a conceptual approach to categorizing cognitive artifacts is perhaps more feasible. A Nersessian-style user-centered approach, however, is particularly suitable for categorizing (cognitive) artifacts deployed for certain cognitive practices in relatively confined boundaries, including scientific research practices in laboratories, navigation practices on a ship, medical practices in a hospital, design practices of architects or engineers in an office, or other cognitive practices in relatively confined environments. Note that I am not saying conceptual and empirical approaches are mutually exclusive.

They are not. Ideally they should cross-fertilize: conceptual approaches should be informed by empirical ones and vice versa, and conceptual approaches should be empirically testable.

2.4 Epistemic Tools

Philosopher of biology Kim Sterelny (2004) refers to artifacts with functional roles in cognitive tasks as "epistemic tools"[7]. He does not give an explicit definition of epistemic tools, but characterizes them by describing five ways we can use such tools. First, by drawing on Clark and Chalmers (1998), Sterelny points out that we alter our environment to ease memory burdens. We store information in diaries, notebooks, and filofaxes so we do not have to remember it ourselves. Second, difficult cognitive problems are transformed into easier perceptual ones. Chess players, for instance, prefer to analyse their moves with a real set and pieces. Trying to figure out your next move and infer its consequences without a real set and pieces in your perceptual field is very hard for most people (see also Kirsh 2009b). Third, difficult perceptual problems are transformed into easier ones. When highlighting text or when labelling an item we decompose our perceptual field into easier detectable elements. Fourth, difficult learning problems are transformed into easier ones. We alter the informational environment of the next generation by spoken and written language, for example when skills are demonstrated such that it is suited for learning purposes. And fifth, workplaces are engineered to make certain tasks easier and more efficient. Sterelny presents research performed by Kirsh (1995) to demonstrate that cooks organise their ingredients and instruments such that they facilitate the cooking process. There is no need to remember the order of ingredients and actions if they are properly structured in your perceptual field. According to Sterelny, these five different ways of using tools show that we interact with and transform our environment for epistemic purposes, thereby creating epistemic tools.

Sterelny's brief taxonomy categorizes epistemic tools on the basis of their function, i.e., the cognitive purpose to which they contribute. So we use notebooks for the purpose of easing memory burdens, label items for the purpose of making a difficult perceptual problem easier, transform our work environment to make particular tasks easier, and so

[7] Sterelny also uses the terms "epistemic artifact" and "tools for thinking" (for the latter term, see also Dennett 2000).

on. It slightly resembles Norman and Brey's cognition-centered approaches in that it takes a cognitive agent as a point of departure, but rather than categorizing tools on the basis of distinct cognitive process to which they contribute, Sterelny takes a *function-centered* approach. But like Brey's taxonomy, Sterelny's taxonomy has nothing particular to say about the specific properties of the artifact, other than that their function is to contribute to realizing cognitive purposes, because it is basically a taxonomy of different types of cognitive purposes, not of artifacts. Sterelny's taxonomy also does not make any specific claims about the informational or metaphysical properties of epistemic tools, although that is not Sterelny's goal.

2.5 Cognitive Technology

A final and quite influential characterization of the functional role of artifacts in performing cognitive tasks comes from Clark (2001). He defends the proposition that a great deal of human cognition is

"rooted in the operation of the same basic kinds of capacity used for on-line, adaptive response, but tuned and applied to the special domain of *external and/or artificial cognitive aids* - the domain of *wideware or cognitive technology*" (Clark 2001, p. 141, original italics)[8].

Cognitive technologies, on Clark's view, are thus external and/or artificial cognitive aids. This is still rather general and Clark does not give an explicit definition of cognitive technology, but (like Sterelny) continues by giving a variety of examples. Some of those examples include simple external memory cues such as leaving a film on your desk as a reminder that it needs developing, or putting a post-it note with the words "develop film" on your computer screen. Other examples concern more complex memory aids such as, for example, Otto and his notebook (Clark & Chalmers 1998). Otto has Alzheimer's disease and therefore has a notebook in which he writes important information (see also Menary 2012). Due to his poor biological memory, he heavily relies on information in his notebook, which is essential for him to successfully get around in the world. Clark & Chalmers (1998) point out four properties that characterize the relation between Otto and his notebook. First, the notebook is a

[8] Clark (2008b) also uses the terms "cognitive artifacts", "material symbols", and "epistemic artifacts" to designate human-made, physical objects with functional roles in cognitive tasks.

constant is Otto's life. When he needs information in the notebook he will rarely take action without consulting it. Second, the information in the notebook is directly and easily available. Third, when retrieving information from the notebook he automatically endorses and the information should furthermore be seen as trustworthy. And fourth, the information in the notebook was believed to be true somewhere in the past and is written down because of this. Given these relational properties, the notebook plays the same functional role as biological memory does for healthy agents and is thus an extended memory and belief system.

The above examples show that we offload information onto our environment, in that way easing our memory burdens. However, offloading information onto our environment does much more than merely easing memory burdens. As Donald and Norman also pointed out, it allows us to perform operations on the offloaded information that are very hard, if not impossible, to perform in the brain. For this reason, Clark is particularly interested in cases in which we think in close interaction with cognitive technology. Consider three of his examples. First, when we write an academic paper, more than just the brain contributes to the writing process. We draw on other people's work, we highlight key concepts and important passages, we use old notes and summaries, we use sketches of arguments, and a word-processer and/or pen and paper. The writing process entails organising and interacting with these elements, these cognitive technologies, such that they effectively result in a well-structured paper.

A second example Clark mentions, drawing on Rumelhart, McClelland, Smolensky & Hinton (1986), is performing a difficult calculation with pen and paper. When we try to perform a calculation (say, multiplying 137 by 363) with pen and paper, the cognitive technologies allow us to carry out manipulations on numerical symbols and store intermediate outcomes externally, thereby breaking the cognitive task into smaller and easier to perform calculations that we already command. For most people, this would be very hard, if not impossible, to do without external aids such a pen, paper and numerical symbols. The idea here is that an embodied brain plus cognitive technologies should be seen as the problem-solving engine, rather than the brain alone.

A third example Clark mentions, drawing on van Leeuwen, Verstijnen and Hekkert (1999), concerns the use of a sketchpad by an artist (see also Tversky et al 2003; Tversky & Suwa 2009). Why does an artist sketch and not just imagine a work of art

internally and then fully offload it onto the sketchpad or canvas? Mental imagery, Clark argues, is constrained in that it is relatively fixed (it is difficult to see new forms and components of a mental image) and limited (there is only so much we can imagine internally)[9]. In order to complement these shortcomings, an artist engages in a process of trial-and-error of sketching, perceiving, re-sketching, perceiving, and so forth, which is an integral element in an ongoing artistic cognitive process. This not only relieves the limited capacity of working memory, but also allows an artist (or designer) to examine, manipulate, prompt new ideas, and revise external information.

In the above three examples, cognitive technology is not the end product of a cognitive process, dangling at the end of a causal chain, but an integrated part of an *ongoing* cognitive process. In cases like these, there is a two-way interaction between human agent and cognitive technology and both components have an active causal role, thereby creating a "coupled system" (Clark and Chalmers 1998). Remove the technological element from the equation and the overall system will drop in behavioural and cognitive competence. In such coupled systems, there is "continuous reciprocal causation" between agent and artifact (Clark 1997, p. 164). Due to this reciprocity, the technology is integrated much deeper into the onboard cognitive system and also has a much stronger transformative impact on internal states and processes. We couple with cognitive technology, Clark argues, to

"reshape and expand the space of human reason. We deploy non-biological wideware (instruments, media, notations) to complement our basic biological modes of processing, creating extended cognitive systems whose computational and problem-solving profiles are quite different from those of the naked brain" (Clark 2001, p. 150).

Thus, cognitive technology expands the space of human cognition, i.e., it makes us more powerful and versatile cognitive agents by complementing the brain's way of information-processing and information-storage. Finally, note that Clark does not try to categorize or systematize cognitive technologies. He usually proceeds by focusing on particular examples, which may be one of the reasons why he called for a systematic taxonomy of different types of external cognitive scaffold and how they aid and hinder

[9] These limitations of mental imagery relate to the shortcoming of engrams that Donald pointed out. Donald argued that engrams are fixed and constrained in their format and capacity, see section 2.1.

cognitive performance (Clark 2002). Clark's call for such a taxonomy was one of the motivations for part I of this thesis. However, in the introduction I argued that we should aim for a broader more inclusive picture that also includes cognitive techniques. A first attempt to paint such a broader picture is developed in the next section.

3. Analysis and Synthesis

In the previous section, I outlined and discussed various notions of artifacts used for performing cognitive tasks. In discussing these notions, I distinguished between cognitive artifacts, cognitive naturefacts, cognitive techniques, and social scaffolds. In this section, I mainly focus on cognitive artifacts and cognitive techniques. I first evaluate the approaches to categorizing cognitive artifacts developed by Donald (1993), Norman (1991), Brey (2005), Nersessian (2003), and Sterelny (2005) (section 3.1). I then clarify cognitive functions of artifacts (section 3.2) and distinguish between cognitive artifacts with representational and non-representational properties (section 3.3). Next, to better understand artifacts with cognitive functions, I briefly contrast them with embodied artifacts which typically (though not necessarily) lack cognitive functions (section 3.4). Thereafter, I further develop the notion of cognitive techniques (section 3.5) by distinguishing between two broad types of such techniques: those that substitute external artifacts and those that concern interactive functional skills (section 3.6).

3.1 Methods of Classifying

Donald loosely classified exogram systems on the basis of their historical properties, i.e., in the historical order in which they have approximately emerged (see table 1). There was no classification based on the functional or informational properties of exograms. Norman argued that a cognitive artifact is categorized purely by its functional role in either experiential or reflective cognition, not by representational properties of the artifact, although these may give some indication of category membership. His two-part classification is therefore basically a classification of different modes of cognition, not of artifacts. Inspired by Norman, Brey categorized artifacts on the basis of the specific cognitive process they aid. Like Norman, Brey takes a human organism and its cognitive processes as a point of departure and then categorizes artifacts based on which cognitive process they aid. Thus, like Norman's

classification, it is basically a classification of different types of cognitive processes, not of artifacts.

Nersessian presented an ethnographic, user-centered, and bottom-up approach to categorizing (cognitive) artifacts in a scientific laboratory, which is different from Donald, Norman and Brey's approaches in which a theorist groups artifacts in a conceptual, top-down manner. A Nersessian-style approach takes artifact-users as a point of departure by asking those users to categorize artifacts on the basis of their experience. Her approach is thus ethnographic in nature and is particularly suitable for categorizing (cognitive) artifacts used for specific cognitive practices in relatively confined boundaries such as a scientific laboratory. Sterelny, finally, takes a purpose-centered approach in that it categorizes artifacts on the basis of the cognitive purpose to which they contribute. It slightly resembles Norman and Brey's cognition-centered approaches in that it takes a cognitive agent as a point of departure, but rather than categorizing artifacts on the basis of the cognitive process to which they contribute, Sterelny takes cognitive purposes as a point of departure. But his classification has nothing particular to say about the specific properties of artifacts, other than that they contribute to realizing cognitive purposes, because it is basically a taxonomy of different types of cognitive purposes, not of artifacts.

These approaches are all important and helpful in that they provide structure and insight in the range and types of artifacts with cognitive functions. However, these approaches have been developed not to better understand artifacts, but to better understand human cognition. For this reason, their classifications are rather short and are typically mentioned in passing. They are usually developed to briefly illustrate that there are indeed different kinds of cognitive artifacts and then continue by focussing on one or two particular examples. Moreover, all these theorists, in one way or another, start with human cognition and then work their way towards the artifacts that agents deploy to aid their cognition and realize their cognitive purposes. These approaches are thus *cognition-centered* in that they take cognition as a point of departure for their analyses. As a result, most of these approaches (with the exception of Donald and Norman) are not particularly interested in or have a great deal to say about the properties of cognitive artifacts, because this is not what they are trying to explain. Most of these approaches also tend to focus on artifacts that are intentionally designed and selected to aid performing cognitive tasks and often (though not always) overlook

artifacts that are improvised. More on this in chapters 3 and 4. In this subsection (and particularly in the next two chapters), I propose an alternative approach to categorizing cognitive artifacts in which I take an *information-centered* approach, i.e., I take as my starting point the specific informational properties of cognitive artifacts and then categorize them on the basis of those properties.

In a sense, the approach I propose here is the opposite of current approaches, as I start with informational properties of artifacts and then I work my way towards better understanding human cognition. Ultimately, my goal is also to better understand human cognition, or, more precisely, situated cognitive systems, I just take a different route to achieving my goal, i.e., from properties of artifacts to cognition, rather than vice versa. My information-centered approach is not to be confused with what Norman (1993) calls a "machine-centered view" of designing, which concerns the design of technologies without taking into account the particular needs, skills, and goals of the users of those technologies, resulting in technologies that do not sufficiently or adequately complement the abilities of their users. Norman and many others argue for a human-centered view on designing, which means that designers should start with understanding the needs, skills, and goals of the users, and then design technologies that fit and complement those attributes. I am very much in favour of this view on designing, but, as I have argued in the previous section, a better understanding of the artifactual element in a situated cognitive system, gives us a better understanding of the epistemic, interactive, and integrative process between skill, artifact, and organism. Focussing on the functional and informational properties of artifacts allows us to see which cognitive capacities and functional skills an agent ought to have to successfully deploy the artifact and may help us in developing a human-centered view on designing cognitive artifacts. Thus, my notion of an information-centered approach to categorizing cognitive artifacts is very different from the notion of machine-centered designing, which I reject.

3.2 Cognitive Functions

I now clarify the cognitive functions of artifacts by analysing and comparing the above outlined notions of exograms as symbolic technologies, cognitive artifacts, epistemic tools, and cognitive technology. These notions combine two elements: (1) an artifactual or technological one and (2) a symbolic, epistemic, or cognitive one. There are thus at least two areas where these notions overlap. First, exograms, cognitive artifacts,

epistemic tools, and cognitive technologies are *human-made, physical objects*. They are thus artifactual in the sense that they are intentionally designed, manufactured, and used for some cognitive purpose by human agents[10]. Second, and more importantly, the above mentioned artifacts are deployed by human agents for the purpose of *functionally contributing to performing a cognitive task*. What these artifacts thus have in common is that they somehow contribute to performing a cognitive task. So, a distinctive property of these artifacts, in relation to other artifacts, is their function, which is cognitive as opposed to pragmatic (see also Kirsh & Maglio 1994; Clark & Chalmers 1998; contrast Loader 2012)[11]. In other words, their function is to provide information that is important to perform a cognitive task, either to make the task easier, faster, more reliable and thus less error-prone, or possible in the first place. More specifically, they are used for performing high-level cognitive tasks, including but not restricted to memorizing (e.g. notebooks), navigating (e.g. maps), planning (e.g. diaries), calculating (e.g. calculators), measuring (e.g. rulers), learning (e.g. textbooks), and artistic cognition (e.g. sketchpads). Thus what the above theorists agree upon is that artifacts with cognitive functions are human-made, physical objects that provide information which is helpful to perform some cognitive task.

One difference in their accounts concerns the degree of malleability of artifacts with cognitive functions. Donald, Norman, and Clark point out that exograms, cognitive artifacts, and cognitive technology can be manipulated and adjusted during a cognitive task. Donald argued that exograms can be arranged and modified in various ways, which is an iterative process that enables reflection and further visual processing. Norman argued that reflective artifacts contain representations that can be modified. Thinking with the aid of such representations, Norman argued, allows us to discover higher-order relationships, structures, or (in)consistencies in the world. So their function is to facilitate and aid reflective, higher-order cognition by providing ways to modify and act

[10] The manufacturing of the artifactual element in a situated cognitive system requires human agency. Although this is an interesting and essential element, the manufacturing of all artifacts requires agency (cf. Hilpinen 1992, 1993, 2011; Baker 2004), so in this regard the above mentioned artifacts are not distinctive.

[11] Although artifacts can have more than just cognitive or pragmatic functions: some may have religious functions (e.g. alters), social functions (e.g. wedding rings), aesthetic functions (e.g. artworks), safety functions (e.g., traffic lights), political functions (e.g. e-voting systems), or moral functions (e.g. speed bumps). For further conceptualizations of moral functions of artifacts, see e.g. Latour (1992) and Verbeek (2011) and for other non-cognitive functions Crilly (2010).

upon representations. Likewise, Clark emphasized the two-way integration of cognitive technology into an ongoing cognitive task. A key notion in Clark's account is the continuous reciprocal causation between agent and technology which creates a coupled system. Although the other theorists certainly do not deny the existence of modifiable representations, they do not make it explicit in their accounts. Conversely, Donald, Norman, and Clark also include into their accounts static artifacts with cognitive functions that do not aid further reflection, but merely provide information about some aspect of the world. A second difference in these accounts concerns whether such artifacts exhibit representational or non-representational properties, which is further discussed in the next section.

3.3 Two Categories of Cognitive Artifacts

Throughout this chapter, I argued for the importance of a category and taxonomy of cognitive artifacts. In the previous subsection, I argued that cognitive artifacts are defined by their function, i.e., their functional role in performing a cognitive task, implying that artifacts with such functional roles are members of the category of cognitive artifacts. In this subsection, I make a first step towards creating such a taxonomy by distinguishing two broad categories of cognitive artifacts, namely, those with representational and non-representational properties.

All of the above theorists recognize that some artifacts have representational properties. However, Donald, Norman, Brey, and Nersessian emphasize that exhibiting representational properties is a necessary condition for having a cognitive function. A cognitive artifact thus has to contain a representation that represents some aspect(s) of the world. In other words, it has to exhibit aboutness or representational content. All the examples that they put forward have representational content. This content is then stored, processed, or otherwise deployed for some cognitive purpose by human agents. It is *because* of their representational content - *i.e. by providing information about some aspect of the world* - that, on their view, artifacts can fulfil cognitive functions. On their view, no representation, no information, and thus no cognitive function[12].

But not all artifacts obtain their cognitive functions through exhibiting representational properties. Otherwise put, not all artifacts aid cognition by providing

[12] In chapter 4, three different types of representational artifacts are distinguished and further conceptualized.

information about some aspect of the world. Donald, Norman, and Clark argued that some artifacts exhibit malleable representational structures that afford adding, erasing, copying, restructuring, reformatting, or otherwise manipulating external representations (see also Menary 2007b; Kirsh 2010). Partly due to their *representational malleability*, these artifacts afford and facilitate higher-order cognition. Consider Clark's examples of writing an academic paper, making a difficult calculation with pen, paper, and numerals, and preliminary sketching. In these cases, the artifactual elements do not have a functional role in performing a cognitive task solely because they have representational properties, but also because they have structural and facilitative properties. Language, numerals, and sketches have content, however, the functional role of the artifacts that help us write papers, perform calculations, and create sketches is at least partly *non-representational*, i.e., to structure the task space and facilitate manipulations of the external structures, thereby contributing to the next step in an ongoing cognitive task. When we write a paper, perform a calculation with pen, paper, and numerals, or sketch, we create an external task space whose malleable structure is equally important as its representational content for performing the task. Those malleable structural properties of the task space are hard, if not impossible, to replicate in an embodied brain, in that way complementing its cognitive capacities. In these cases, both representational and structural properties contribute to performing the task.

There are, however, also cases in which we interact with artifacts that have functional roles in performing cognitive tasks not by exhibiting representational properties, but solely by exhibiting malleable structural properties. Consider Hutchins' and Sterelny's examples: drawing on Kirsh's (1995) notion of "the intelligent use of space", they argued that we sometimes structure our environment such that it facilitates a particular procedure. When cooking, for instance, we organize and transform our task environment such that it makes certain procedures more efficient and easier to perform. We do so by structuring the utensils and ingredients such that it facilitates the cooking process. Structuring utensils and ingredients in a particular way does not have any obvious representational properties. Kirsh (1995) himself makes a tripartite distinction between spatial arrangements that simplify choice, perception, and internal computation. Thus the cognitive function of such spatial arrangements is not representational, but to simplify choice, perception, or internal computation. A distinction may therefore be made between cognitive artifacts exhibiting

representational and non-representational properties. These distinct informational properties are essential for better understanding the range and variety of types of cognitive artifacts and are therefore further conceptualized in chapters 3 and 4.

3.4 Embodied Artifacts

To better understand artifacts with cognitive functions, it is helpful to compare them with artifacts have no obvious functional role in performing cognitive tasks. Coffee cups, chairs, tables, flower pots, light switches, windows, and trash bins, for example, are not used for cognitive purposes, at least not in their normal use, because they have no functional role in performing a cognitive task. Moving a chair so that you can sit on it, or turning on a light switch, does not straightforwardly contribute to performing a cognitive task. We interact with those artifacts to create a change in the state or location of the artifact, not because it aids our cognition, but because that goal state is desirable for some practical or pragmatic purpose. The purpose for which such artifacts are deployed is therefore not cognitive but pragmatic.

Artifacts with pragmatic functions may, of course, influence human cognition, because sometimes we have to think about how to interact with those artifacts. However, being the object *of* perception and cognition is necessary but not sufficient for functionally contributing *to* a cognitive task. Many things are the object of perception and cognition, but do not in any obvious sense help us to perform or complete a cognitive task. But, this is not to say that such artifacts can never have cognitive functions. For example, when I leave a rented DVD on my desk to remind myself to bring it back to the video store, the DVD (in virtue of its location) arguably functions as a mnemonic aid. So during improvisation, we can attribute cognitive functions to initially non-cognitive artifacts, thereby demonstrating that we cognitively exploit not only our natural environment, but also our artifactual environment quite opportunistically. This example shows that a cognitive artifact is neither defined by intrinsic properties of the artifact nor by the intentions of the designer, but by its function, which is established by the intentions of the user and by how it is used.

Other artifacts bear a close relationship to the "body schema", which is a non-conscious neural representation of the body's position and its capabilities for action (Johnson-Frey 2003; Maravita & Iriki 2004; Gallagher 2005). For developmental and evolutionary reasons, body schemas are flexible and fluid and therefore artifacts like hammers,

cricket bats, pencils, spoons, walking canes, and screwdrivers, can be incorporated into the body schema, as to enlarge and extend the representation of the body's position and options for action. Such artifacts are then said to be "embodied" and are not experienced as objects in the environment, but as part of the human motor or perceptual system. Following Brey (2000), two types of embodied artifacts may be distinguished[13]. First, embodied artifacts used to act on or physically change the world such as hammers, screwdrivers, and spoons. Second, embodied artifacts used to better perceive the world such as glasses, binoculars, stethoscopes, or a blind man's cane.

When using embodied artifacts to act on or change the world, an agent does not first intend to act on the artifact and then on the environment. It is not a two-step process. Rather, an agent merely intends to act on the environment through the artifact and does not consciously experience the artifact when doing so (Clark 2007). The perceptual focal point is thus at the artifact–environment interface, rather than at the agent–artifact interface. In this sense, embodied artifacts are transparent (Ihde 1990), or, in Heideggerian terms, ready-to-hand (Heidegger 1962). Such artifacts thus bear a close relationship to the body schema, which is a cognitive construct, and in that sense have a stronger and more obvious relationship to cognition as compared to artifacts such as light switches, chairs, and windows, which are not incorporated into the body schema. However, the purposes for which such embodied artifacts are used are pragmatic rather than epistemic. Hammering a nail into a wall, using a screwdriver to open a can of paint, or a spoon to eat soup, does not directly and straightforwardly contribute to performing a cognitive task. Although hammering, opening a can of paint, and eating are tasks that require cognition, the purposes of those tasks is to physically change the environment with the aid of the artifact because the change they induce is desirable for its own sake. Hammers, screwdrivers, spoons and other embodied artifacts that are used to act on the world are thus typically not utilized for cognitive purposes but for pragmatic purposes.

However, a subgroup of embodied artifacts plays an important facilitative-functional role in performing cognitive tasks. A computer mouse is an artifact deployed to interact with a computer, for example to select a piece of text, open a folder, or delete a document. The mouse is (ideally) transparent in use and withdraws from attention, as

[13] For an analysis of embodiment in relation to virtuality, see Hubert Dreyfus (2001).

we tend to focus on the cursor rather than on the mouse. The mouse itself does not have any epistemic or informational properties, but it does facilitate a relationship to a rather powerful cognitive artifact (i.e. a computer) and is in that sense important for interacting with and creating external representations. Likewise, pens, styluses, and other writing and drawing tools are embodied artifacts which are used to act on the world, but do not have any epistemic properties in themselves. They do, however, allow us to write or draw symbols and structures, thereby facilitating a relationship to and creating notational systems.

Embodied artifacts in the second class are used to better perceive the world. Glasses, binoculars, stethoscopes, or a blind man's cane, are not used to create physical changes in the environment but are used to perceive parts of the environment that are difficult or impossible to perceive without the aid of such artifacts. Particularly interesting in the context of this thesis is a blind man's cane (Merleau-Ponty 1964). For the blind man, the cane is not an external object with which he interacts, but he interacts with the environment through the cane. The focus is on the cane-environment interface, rather than on the agent-cane interface. The cane is furthermore incorporated into his body schema and is experienced as a transparent extension of his motor system. So it bears all the characteristics of an embodied artifact and indeed has become the textbook example of an embodied artifact. But it also has epistemic properties, I claim, because it provides the blind man with tactile and auditory information about the kind and location of objects and structures in his environment. On the basis of this feedback, a blind man can identify the location and kind of objects he is interacting with, which is essential to successfully navigate through the environment. The cane has therefore a quite obvious functional role in performing a cognitive task in that it provides a blind person with information about objects and structures in its environment. The cane is thus used for cognitive purposes and so embodied and cognitive artifacts are not mutually exclusive categories, but sometimes overlap.

When artifacts are incorporated into the body schema, our (intentional) stance towards the world is changed. This is so because such artifacts enable us to perform actions and perceive things that we would otherwise not be able to do or perceive, i.e., our action and perceptual repertoires are enhanced. This change in perception and stance towards the world has been referred as a revealing/concealing structure (Heidegger 1962) or a magnification/reduction structure (Ihde 1990). For example, when observing the night

sky with a telescope, stars are magnified but also taken out of their larger context. So some visual aspects are magnified, while others are reduced (see also Kiran and Verbeek 2010; compare Heersmink 2012a). So the particular properties and affordances (Gibson 1979) of an embodied artifact change our stance towards the world. Further, the skills needed to use such artifacts are also important. When I first start using a hammer, my skills are underdeveloped and the hammer is not yet transparent. But gradually my hammer-using skills develop and the artifact becomes transparent which will then alter my stance towards the world. A hammer affords more options for a skilled hammer-user, as compared to a less skilled hammer-user. In chapter 5, transparency is further developed as one of the dimensions along which to study agent-artifact integration.

Some cognitive artifacts are embodied. Consider, for example, a carpenter who measures the length of some object with the aid of a ruler, a typical representational artifact. When deploying the ruler, the perceptual focal point is not at the agent-ruler interface, but at the ruler-object interface. For an experienced carpenter, the ruler is transparent and incorporated into the body schema. Similarly, in his study of ship navigation, Hutchins described a device called the alidade, which is a telescopic sighting device with a built-in gyrocompass. The scale of the gyrocompass is superimposed on whatever is seen through the alidade and allows its user to indicate bearings relative to the ships head. The alidade is used to better perceive the world and to measure and quantify the amount of degrees of the ship's head in relation to some landmark. For experienced users, the alidade is transparent and incorporated into the body schema of its user, as the perceptual focus is not on the alidade, but on what is seen through it. Further, the revealing/concealing or magnification/reduction structure of embodied tools also applies to embodied cognitive artifacts such as rulers and alidades. When using these artifacts, our perception is geared towards measuring specific properties of the world, while other properties are overlooked or reduced.

Finally, embodiment and body schemas are generally important for interaction with most cognitive artifacts (see also Kirsh 2013). Calculators, computers, diaries, maps, books, and many other external cognitive artifacts are interacted with bodily and although they may not be incorporated into the body schema, in order to successfully interact with the artifact it is important that the body schema effectively guides motor processes that are required for interacting with such artifacts. Thus body schemas can

be seen as the neurological foundation for interactive functional skills and embodied interactions with artifacts and other environmental structures.

3.5 Internalized Information and Cognitive Techniques

Having analysed cognitive functions, distinguished between two categories of cognitive artifacts, and embodied artifacts, I now further develop the notion of a cognitive technique. We have seen that Norman, Hutchins, and Sutton developed a broad notion of cognitive artifacts, as they included what I referred to as internalized information and cognitive techniques into their notion of cognitive artifacts. Norman briefly developed the notion of a "mental artifact". He argued that reading, arithmetic, logic, and language are psychologically efficacious not because of their physical properties, but because of the rules and structures they propose. Although such mental artifacts are physical in that they supervene on neural and bodily structures, they do not obtain their power from physical properties but from information-structures. Other mental artifacts are procedures and routines like mnemonics for remembering or methods for performing certain tasks. Norman argued that mental artifacts are artificial in that they would not exist without human invention. On his view, anything invented by humans for the purpose of improving thought or action should be seen as an artifact, whether it is physical or mental is irrelevant.

Hutchins pointed out that if we focus too much on external artifacts in performing cognitive tasks, it may discourage us to overlook the functional roles of "internal artifacts" and "external structures that are not designed". To support his claim, he argued that Micronesian navigators use a particular perceptual strategy to recognize a certain cluster of stars and to attribute a navigational function to those stars. Such learned perceptual strategies are what Hutchins referred to as internal artifacts. Sutton, finally, developed the notion of an "internalized cognitive artifact". He described how internalized surrogates such as words, phrases, and other mental representations can be co-opted to function as external exograms usually do, for example when we learn a mnemonic or use the method of loci to help us remember.

Norman, Hutchins, and Sutton use the term "artifact" for cognitive states and processes that take place within a human organism. In my discussion of Hutchins' account, I claimed that it is conceptually clarifying to refer to learned perceptual strategies not as internal artifacts, but as cognitive techniques. I argued that artifacts are physical

objects intentionally designed and manufactured for a particular purpose, whereas a technique is a method or procedure for doing something. Because such perceptual strategies are not manufactured physical objects but ways of perceiving, they can best be seen as cognitive techniques. As I think it is important to single out and develop a vocabulary for distinct processes and objects that produce cognitive effects, I believe the word "technique" is more appropriate here as it emphasises the *artificiality* of such cognitive capacities but not the *artifactuality*. This critique also applies to Norman and Sutton's use of the term artifact. Using inductive or deductive logic to make an inference, or a mnemonic or method of loci to remember, can better be seen as cognitive techniques, rather than as mental artifacts or internalized cognitive artifacts[14]. Note that, contrary to Norman, who argued that our real cognitive powers come from external artifacts, I am not privileging external artifacts in any way. I am just saying that the term "artifact" should be reserved for physical objects intentionally designed and manufactured for a particular purpose and that what the above theorists have referred to as mental artifacts, internal artifacts, and internalized cognitive artifacts do not fit those conditions and are thus not proper artifacts.

Furthermore, in my discussion of Sutton's notion of internalized cognitive artifacts, I argued that cognitive artifacts (words, phrases) or artifactual environment (virtual architectures) are not literally internalized, only the information they contain is internalized and soaked-up. So rather than referring to such phenomena as internalized cognitive artifacts, I suggested referring to them as internalized information. This critique also applies to Norman's and Hutchins' accounts. When we learn to read, do arithmetic and logic, or learn to navigate on the basis of the stars, we internalize information. The nature of that information may be different: learning inductive logic, perceptual strategies, mnemonics, or learning to read or do arithmetic concern the internalization of very different types of information. Some concern the internalization of abstract rules, others concern the internalization of language or other kinds of

[14] We may still use the term "internalized cognitive artifact" for manufactured, physical objects that are implanted in the human body that aid us in performing cognitive tasks such as invasive brain-computer interfaces (BCIs) that help paralysed subjects to spell words and sentences. However, note that current BCI systems are only partly internalized and are thus still largely external to the human organism, as only the electrode arrays are implanted into the brain (Heersmink 2011, 2013a). As far as I am aware, there is currently no cognitive artifact that is fully implanted into the human body or brain, although this may of course change in the future.

representational systems, and yet others concern the internalization of non-representational structures. What matters is to recognize that not the artifacts themselves are internalized, but the information they contain.

In a somewhat similar vein as Sutton, Helen de Cruz (2008) presents an interesting example of internalized information. In Chinese education, pupils learn to perform calculations with the aid of an abacus. After a certain period of training, Chinese abacus-users no longer need the physical abacus to perform calculations, as they have internalized the information-structure (i.e. rows of beads) of the abacus. They can then mentally visualize and manipulate abacus beads in order to perform calculations. The Japanese use a similar artifact called a soroban, which has up to sixteen rows of beads and can thus be used to perform very large calculations. Some users have fully internalized the information-structure of a soroban and can thus perform very large calculations with a remarkable accuracy by mentally simulating and manipulating the information-structure of a soroban. The internalized information-structure has the same functional role in performing a cognitive task as a similar external version, i.e., both the external and internalized version of the abacus or soroban have the same functional role in performing calculations.

Thus, learning a cognitive technique is internalizing an information-structure, either from other human beings, from cognitive artifacts, from artifactual environments, or a combination of these. However, cognitive techniques or any other technique need not necessarily be learned from other humans or artifactual structures, they may also be self-taught. When such techniques are learned from other human agents, it may be vertical (i.e. from parent to offspring), oblique (i.e. from the parental to the offspring generation, e.g. from teacher to student), or horizontal (i.e. within-generation transmission, e.g. from student to student) (Cavalli-Sforza & Feldman 1981; Laland 2002). However, quite often the transmission of cognitive techniques from humans to humans is a combination of these three types of information transfer (see also Sterelny 2003, 2012). After the information-structure is firmly internalized and the technique is mastered, it has transformed our mind and (ideally) augmented our cognitive capacities, as we can now perform more cognitive tasks, including navigating on the basis of the stars, recalling the order of the planets in our Solar System, recalling the order of items on a shopping-list, or performing calculations by mentally manipulating

abacus beads. Such techniques are used to perform cognitive tasks (navigating, remembering, calculating) and are thus correctly referred to as *cognitive* techniques.

3.6 Substitution and Functional Skills

It is helpful to distinguish between different types of cognitive techniques and to clarify their relationship to external artifacts and structures. I identify two broad categories of cognitive techniques. The first category of techniques concerns the internalization of information that *substitutes* a cognitive artifact or artifactual structure(s) in the environment. Examples include: internalizing information of mnemonics, abaci, sorobans, and methods of loci. Another example would be the internalization of the information of a rather simple map, for example a map of the layout of the Museum of Contemporary Art (MCA) in Sydney. The MCA has four levels and so the map depicts four simple representations of the layout of each level, indicating with bright colours which exhibition is going on in which room. One may of course use the actual map to find one's way in the museum, but it is fairly easy to memorize the layout of at least one level which typically has three or four main rooms, thereby internalizing an information-structure that (partly) substitutes a cognitive artifact, resulting in a cognitive map. In the above cases, information is internalized and creates cognitive techniques that can be used to remember, calculate, or navigate and the external artifact or structure has then become superfluous. However, as Donald pointed out, biological memories are vulnerable to deterioration and degradation and, consequently, we may sometimes forget our cognitive techniques and then have to relearn and re-internalize them.

The second category of techniques concerns the internalization of information that creates functional interactive skills, allowing us to exploit the environment for cognitive purposes. For example, Micronesian navigators have internalized information that creates a cognitive technique, allowing them to use the stars as navigational aids. A human agent, the cognitive technique, and the environment are needed to perform a cognitive task. Thus agent, technique, and environment constitute a situated cognitive system, rather than merely an agent and technique, as is the case with techniques that substitute a cognitive artifact or artifactual environment.

This type of cognitive technique, however, is much more general and ubiquitous than just this specific example. It also includes our learned ability to interpret external

representational systems such as written language, numerals, diagrams, maps, traffic signs, et cetera, which are all learned perceptual strategies. During ontogenetic development, most people learn how to interpret meaningful information and to use it for their cognitive purposes. Extracting meaningful information from one's environment requires interpretative skills as well as the direct perceptual intake of external representations or other meaningful information. The ability to interpret meaningful information is closely related to and often interdependent with the ability to create meaningful information. For example, reading and writing are deeply interdependent, but are nonetheless different cognitive techniques. Similarly, an architect designing a blueprint for a building and a construction worker using it to help build the actual building are employing two different techniques. However, both the ability to interpret and create external information result in functional interactive skills that allows us to exploit the environment for cognitive purposes. These functional interactive skills are needed to interact with cognitive artifacts and are thus constitutive elements of situated cognitive systems. Thus, human agents internalize a variety of information, resulting in a two broad categories of cognitive techniques, those that substitute cognitive artifacts and those that concern interactive functional skills that allow us to exploit our environment for cognitive purposes.

4. Concluding Summary

This chapter began by reviewing and discussing several characterizations and classifications of artifacts that aid us in performing cognitive tasks. I discussed the notions of exograms (Donald), cognitive artifacts (Norman, Brey, Hutchins, Nersessian), epistemic tools (Sterelny), and cognitive technology (Clark) and pointed out that the defining property of such artifacts is their functional role in performing a cognitive task, i.e., they are used for used cognitive functions, as opposed to pragmatic functions. More specifically, their function is to provide task-relevant information, in that way making certain cognitive tasks easier, faster, more reliable, or possible at all. In order to better understand the distinctiveness of artifacts exhibiting cognitive functions, I contrasted cognitive functions with pragmatic functions by building on and further developing the notion of an embodied tool, which typically (though not necessarily) lacks cognitive functions. Finally, I distinguished between two broad categories of cognitive artifacts, those with representational and ecological properties.

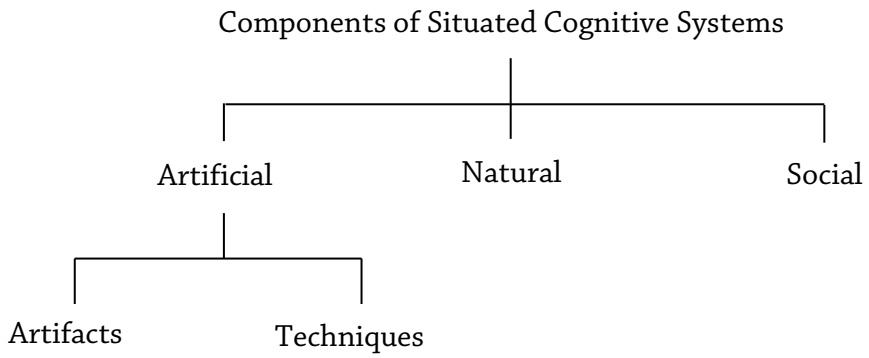

Figure 1. A classification of the components of situated cognitive systems.

In case of techniques, I discussed the notions of mental artifacts (Norman), internal artifacts (Hutchins), and internalized cognitive artifacts (Sutton). On the basis of those notions, I developed the concept of internalized information and distinguished between two broad categories of cognitive techniques, those that substitute cognitive artifacts and those that concern interactive functional skills that allow us to exploit our environment for cognitive purposes. In addition to human-made or artificial scaffolds - i.e., artifacts and techniques - we also deploy cognitive naturefacts and other people to help us perform our cognitive tasks. Thus the components of situated cognitive systems may consist of artificial (i.e., artifacts and techniques), natural, and social scaffolds (see figure 1).

3

Cognitive Artifacts, Functions, and Classification Systems

1. Introduction

In the previous chapter, I outlined an initial classification of the components of situated cognitive systems, including those that are artificial, natural, and social. The goal of this chapter is to provide the tools to further conceptualize the category of cognitive artifacts. In the next chapter, I develop a detailed taxonomy in which cognitive artifacts with similar informational properties can be grouped into categories. But before I do this, I need to demarcate the target domain and provide some of the tools to develop a method for taxonomizing cognitive artifacts. We have seen that the functional role of token artifacts in performing cognitive tasks has been widely discussed in the philosophical and cognitive science literatures under the genus of situated cognition and its species of embodied, embedded, extended, and distributed cognition (Norman 1993, Hutchins 1995, Clancey 1997; Clark 1997, 2003, 2008b, Donald 1991, 2010; Rowlands 1999, 2010; Dourish 2001; Anderson 2003; Wilson 2004; Dascal & Dror 2005; Zhang & Patel 2006; Suchman 2007; Robbins & Aydede 2009a).

However, metaphysical issues concerning (the categories of) cognitive artifacts have not received much attention, because the explanatory targets of the situated cognition movement are situated cognitive *systems*, not cognitive *artifacts*. Whilst this explanatory focus is understandable, a better understanding of the artifactual element in situated cognitive systems will result in a better understanding of the overall situated system, which is addressed in Part II of this thesis. Furthermore, current classifications as discussed in the previous chapter are cognition-centered in that they start with human cognition and then work their way towards the representational artifacts that agents deploy to aid their cognition and realize their cognitive purposes. Consequently, most of these approaches (with Norman and Donald as an exception) do not have a great deal to say about the particular metaphysical and informational properties and categories of cognitive artifacts, because that is not their explanatory target.

While other authors working in philosophy of technology have started to address metaphysical issues regarding technological artifacts (Preston 1998b, 2009, 2013; Kroes & Meijers 2001; Meijers 2001; Bakker 2004; Lawson 2008; Thomasson 2009; Houkes & Vermaas 2009, 2010) and to a lesser extent category issues (Carrara & Mingardo 2013; Houkes & Vermaas 2013; Franssen, Kroes, Reydon & Vermaas 2014), they focus on artifacts in general and have little, if anything, particular to say about cognitive artifacts. So there is a gap in the literature which this chapter and the next address by conceptualizing metaphysical and category issues pertaining to cognitive artifacts. By doing so, these chapters fill a gap in the literature and bring into contact concepts and theories in (philosophy of) cognitive science and philosophy of technology and strengthen the rather thin ties between those fields, which is beneficial for both fields[15]. The research in these chapters thus builds intradisciplinary and interdisciplinary bridges.

Developing a systematic taxonomy of the artifactual element in situated cognitive systems is important for at least two other reasons. First, because a substantial part of

[15] Preston (2014) argues along similar lines. She argues that the metaphysics of artifacts is somewhat isolated from other academic disciplines and would benefit from contact with anthropology and archaeology, as these disciplines have studied artifacts extensively. This chapter and the next is an attempt to build such connections.

our cognitive activity quite heavily involves the use of artifacts, perhaps more so than other scaffolds, it is vital to have a taxonomy that gives us a richer and deeper understanding of the many distinct kinds of cognitive artifacts and their informational properties. If we are in essence "natural-born cyborgs", as Clark (2003) has claimed, then it is important to better understand the rich variety of relationships that are established between us and different kinds of artifacts that make us natural-born cyborgs. Only when we have a proper understanding of the functional and informational properties of such artifacts, are we in a position to address and better understand the numerous relationships that are established to such artifacts. This is important because it provides us with a better and deeper understanding of cognitive agents as tool-users, which is generally considered as one of the distinctive features of human agents (Vaesen 2012).

Second, it contributes to the general project of the metaphysics of artifacts. This is a young and exciting, but unfortunately a much underdeveloped subfield in the philosophy of technology, which has traditionally focused on social, cultural, and ethical consequences of technology on human beings and society at large (see for overviews Achterhuis 2001; Verbeek 2005; Dusek 2006; Brey 2010; Reydon 2012). Carl Mitcham (1994) has referred to this type of philosophy as "humanities philosophy of technology". However, in the last decade or so, a new branch of philosophy of technology is emerging that is sometimes referred to as philosophy of artifacts or "internal philosophy of technology", which is much more concerned with technological artifacts in themselves (Franssen, Lokhorst & van de Poel 2009). So rather than focusing on social, cultural, and ethical issues, philosophy of artifacts focuses more on methodological, epistemological, and metaphysical issues concerning technology (Kroes & Meijers 2001). This chapter and the next contribute to expanding and further developing of the philosophy of artifacts by analysing and conceptualizing metaphysical, informational, and functional properties of cognitive artifacts.

I proceed as follows. The chapter starts with identifying and demarcating the target domain by conceptualizing cognitive artifacts as a functional kind, i.e., a kind of artifact that is purely defined by its function. More specifically, I develop a pluralist notion of functional kind by building on Preston's (1998b, 2013) notions of proper function and system function. Having identified the target domain, it then briefly looks at the multiple usability of physical structures, the multiple realizability of cognitive function,

and malfunction (section 2). Next, I look at classification systems in archaeology (morphological, historical, and functional classifications) (section 3). In the next chapter, these classification systems are synthesized into a method for taxonomizing cognitive artifacts. So first I demarcate the target domain, then present the theoretical tools to taxonomize that target domain, and thereafter (in the next chapter) I use those tools to develop a taxonomy of cognitive artifacts. Finally, a concluding summary is given (section 4).

2. Identifying the Target Domain

In the previous chapter (section 3.2), I argued that a distinctive and category defining property of cognitive artifacts is their functional role in performing a cognitive task, i.e., they are used for cognitive purposes, whereas most other artifacts are used for pragmatic purposes. So what distinguishes them from other artifacts is the purpose for which they are used, i.e., their *function*. More specifically, their function is to provide representational or ecological information, thereby making certain cognitive tasks easier, faster, more reliable, or possible at all. Thus their most important property, I claim, is their function (compare Thomasson 2007).

Like all artifacts, cognitive artifacts have many other properties, e.g., those pertaining to physical, design, manufacturing, and historical characteristics, but also those pertaining to issues such as style, quality, and price. An abacus, for example, may have ten or fifteen rows of beads, it may have wooden or plastic beads, it may be black or red, manufactured by an artisan or by mass production, it may have been manufactured in ancient Rome or modern Sydney, its style may be fashionable or old-fashioned, it may be of poor quality or high quality, or it may be expensive or inexpensive. All these properties are of course relevant for *characterizing*[16] an abacus, but what *defines* it is its function, which is to aid its user in performing calculations: a paradigm cognitive task. I therefore demarcate the boundaries of the target domain by conceptualizing cognitive artifacts as a functional kind, i.e., a kind of artifact that is purely defined by its function[17] (Franssen, Lokhorst & van der Poel 2009; see also Carrara & Vermaas 2009).

[16] Non-functional properties such as these are still important for characterizing the intentional aspect of artifacts, see Vaesen (2011b).

[17] Building on this notion of a functional kind, one may identify other functional kinds such as artifacts that have as their function to transport (e.g. cars, busses, trains, bicycles, airplanes), to contain (e.g. bottles,

In Hilary Kornblith's words: "At least for the most part, it seems that what makes two artifacts members of the same kind is that they perform the same function" (Kornblith 1980, p. 112).

2.1 Artifact Functions

Given the centrality of function in the demarcation of the artifactual element in situated cognitive systems from other artifacts, it is important to further explain what a function is and to distinguish between two kinds of function: proper and system function. In the next subsection, I use these two kinds of function to illustrate which artifacts can be considered as members of the functional kind of cognitive artifacts.

The contemporary debate on function in philosophy of science originated in the works of Carl Gustav Hempel and Ernest Nagel, who were concerned with functional explanation in science, particularly in biology. Larry Wright, Robert Cummins, Ruth Garrett Millikan, Karen Neander, and others later added substantial content to the function debate in philosophy of biology, which remains a prominent topic of discussion. Philosophy of mind and philosophy of cognitive science also have given a fair share of attention to the notion of function. In functionalism, mental states are conceptualized not by their material constitution, but according to their functional role in an overall mental economy. Thus the nature of a mental state is identified by its causal-functional role in relation to sensory input, other mental states, and behavioural output (Putnam 1975; Levin 2009).

However, relatively little attention has been given to the concept of function in philosophy of technology. Only in the last decade or so, have philosophers tried to articulate and clarify a function theory for artifacts (Preston 1998b, 2010, 2013; Kroes 2006, 2012; Meijers 2001; Houkes & Vermaas 2004, 2010; Scheele 2006; Margolis & Laurence 2007). One of the first philosophers who has given a substantial account of function for the artifactual domain is Preston (1998b). Building on Wright (1974), Cummins (1975), Millikan (1984), and others, she develops a pluralist theory of function for artifacts that combines the notions of proper function and system

---

drawers, cups, trash cans, pen trays, boxes, bags), or to provide light (e.g. torches, oil lamps, candles, lanterns, fluorescent lamps, incandescent lamps). There are, likewise, many different sorts of functional kinds. Moreover, a token artifact can be a member of more than one functional kind, as some artifacts have more than one function.

function. Some philosophers have tried to reduce one of those notions to the other, arguing that there ought to be a universal account of function. But Preston argues that both notions explain different phenomena and both are needed for a function theory of artifacts. Let me briefly explain these two kinds of function.

2.1.1 Proper Functions

Millikan (1984, 1989; see also Griffiths 1993) has argued for an etiological theory of function-ascription, which argues that in order to understand the function of a biological trait or technological artifact one has to take into account the causal-historical background of that trait or artifact. She develops the concept of proper function which gets established by a causal selection history. For example, hearts pump blood and in doing so they also make a certain sound, which has nothing to do with their function. Hearts do not exist because they make a certain sound, but because their function (i.e. pumping blood) has contributed to the fitness of the organism. The sounds that hearts make are a mere epiphenomenon of the mechanical workings of the heart and are irrelevant for the successful reproduction of the organism and thus the organ. Natural selection does not and cannot select on the basis of epiphenomena, but on the basis of effectively performing functions that contribute to successful reproduction. Likewise, chairs exist and are re-produced because they are widely used to sit on, not because you can stand on them or hang your coat on them. So if one wants to understand the function of an organ, trait, or artifact, one has to look at the selection history of that entity. For the biological domain, this is an evolutionary history of natural selection. For the artifactual domain, it is a history of cultural selection by users.

Thus, proper functions of artifacts are established through a process of cultural selection. Artifacts are designed or invented, and if they are successful in performing their function, they will be re-produced. Somewhere in the past, chairs have been invented to sit on and have been quite successful in their re-production. So the proper function of a chair is to sit on, because they have been selected for this purpose by previous generations of users. Chairs can also be used for other purposes, for example to hang your coat on, to stand on, or to block a door from opening. But these purposes are not the reason why chairs are selected by their users and are therefore not their proper function. Most artifacts have one proper function: chairs are for sitting, pens are for writing, coffee cups are for containing coffee, magnifying glasses are for magnifying

objects or structures, cars are for transportation, et cetera. Some artifacts have more than one proper function. Certain pencils have an eraser at their end, so that one can use it for both writing or drawing and erasing. Yet other artifacts have many proper functions. A Swiss Army knife, for example, is for cutting, opening bottles, opening cans, sawing materials, clipping nails, and so on. Swiss Army knives are selected by their users not for one particular function but for many functions and are thus a multifunctional artifact with numerous proper functions. So a number of proper functions can coexist in one artifact.

2.1.2 System Functions

Cummins (1975, 1983) argued that functional explanation in science is not based on causal-historical selection, but on the current capacities and dispositions of a whole system in terms of its components. By giving a number of counterexamples, Cummins argues that causal-historical selection cannot explain the existence of certain organs or traits. Likewise, in the artifactual domain there are numerous examples of artifacts that are used for purposes that were not intended by their designers. Chairs are used to stand on, old tires are used as swings, screwdrivers are used for openings cans of paint, et cetera. In such instances, the function of those artifacts has nothing to do with their history of cultural selection and has everything to do with their current capacities or dispositions in a given context. Thus, the current function of an organ, trait, or artifact is not necessarily linked to its selection history, but is in principle divergent. Preston (1998b, 2013) refers to Cummins' notion of function as "system function" and I will use her terminology throughout this thesis[18].

System functions of artifacts are either improvised uses of artifacts or the functions of novel prototypes. In case of novel prototypes, the first generations only have system functions, which are over time consolidated into proper functions. In case of improvised uses of artifacts, Preston develops two conceptions of system function in analogy to exaptations in biology. An exaptation occurs when a biological trait evolves such that it loses its original function and obtains a new function. Bird feathers are a classic example, which initially evolved for insulation, but were later adapted for aerodynamic purposes during flight. Preston makes a distinction between two types of

[18] Cummins' notion of function has also been referred to as causal role function (Neander 1991; Houkes & Vermaas 2010) because the function of some entity is determined by its causal role in a larger system.

system function, namely, "standardized ongoing exaptations" and "idiosyncratic ongoing exaptations". Standardized ongoing exaptations are repeated uses of artifacts that are not their proper function. Examples include using a chair to stand on, a screwdriver to open a can of paint, or a spoon to open a cocoa tin. Such uses are not intended by their designers and are not the reason why such artifacts are selected by their users, but they are nevertheless widespread cultural practice. Such system functions are quite often ongoing additions to the artifacts' proper function. Hence, proper functions and system functions can coexist in one artifact.

Idiosyncratic ongoing exaptations are improvised uses of artifacts by individuals. They are not widespread cultural practice, but ongoing uses of artifacts for individuals or small groups of individuals. Preston puts forward a number of examples, including the use of an old cast-iron as a bookend, or using a shoe-string to tie up a tomato plant. Such idiosyncratic uses of artifacts may over time become more established and may spread to other social groups. They could potentially even become a consolidated proper function. Preston argues that the notion of system function is "crucial in understanding the history of hominid tool use, which developed from the simple exaptive use of naturally occurring objects as hammers, digging sticks, and so on, to the pervasively artifactual environment we Western industrialized humans inhabit today" (Preston 1998b, p. 253). System functions are thus important because they explain the development of artifacts and technology. A function theory including merely proper functions would only be able to account for certain uses of artifacts and not for others. Proper and system function seem to explain different phenomena, which is an argument for a pluralist theory of function.

I would like to add a third type of system function, namely, "idiosyncratic exaptations". Preston's examples of standardized and idiosyncratic exaptations are *ongoing*. Standardized ongoing exaptations are often repeated uses of an artifact by a community of users and idiosyncratic ongoing exaptations are ongoing uses of artifacts for individuals or small groups of individuals. So the former are culturally well-established uses of artifacts for a community and the latter are well-established (or fairly well-established) uses of artifacts for (small groups of) individuals. However, there are also cases of idiosyncratic uses of artifacts that are not well-established but one-offs, which are neither culturally widespread nor are they ongoing. For instance, someone may use a screwdriver as a weapon in order to defend oneself, or someone might use a hammer

to break a window to get into the house because she has forgotten her key. Such uses can be called idiosyncratic exaptations, as they may occur only once or twice in a lifetime.

Finally, the above distinctions should perhaps not be seen as strict subcategories of system function, but rather as points on a continuum of system functions. Proper functions are rather constant over time, but system functions come and go. They may vary between improvised one-offs, ongoing and well-established uses of artifacts for individuals, or widespread uses of artifacts for a community of users, and everything in between. There are no clear-cut criteria to indicate when idiosyncratic exaptations become idiosyncratic ongoing exaptations. There are, likewise, no clear-cut criteria to indicate when idiosyncratic ongoing exaptations become standardized ongoing exaptations. It is therefore better to conceive of system functions as a continuum.

2.2 A Pluralist Notion of Functional Kind

Through using Preston's concepts of proper and system function, I now further specify the target domain by developing a pluralist notion of the functional kind of cognitive artifacts. It is pluralist in the sense that it contains cognitive artifacts with both proper and system functions. Those with proper functions (i.e. proper cognitive artifacts) have a history of cultural selection, whereas those with system functions (i.e. system cognitive artifacts) are improvised uses of initially non-cognitive artifacts. Abacuses, for example, have a long history of cultural selection (and thus re-production), as they have been selected by their users at least since the invention of the Salamis Tablet, an abacus-like device that was used for performing calculations, dating back to roughly 300 B.C. The proper function of an abacus is therefore to aid its user in performing a calculations, because it has been designed and selected for that purpose. Abacuses may also be used for other purposes, e.g., as a kids' toy, but this is not the reason why they are designed and not the reason why they are selected by their users and is thus not their proper function.

Likewise, maps, computers, calendars, radars, rulers, encyclopaedias, and countless other cognitive artifacts are designed to perform cognitive functions and have a history of cultural selection and thus have as their proper function to aid their users in performing cognitive tasks. It may therefore be argued that such artifacts have a *cognitive proper function*. Some of these artifacts have more than one cognitive proper

function. Computers, for instance, are highly multifunctional devices and may be seen as the Swiss Army knife of the cognitive artifacts (see also Norman 1998; Hollan, Hutchins & Kirsh 2000; Brey & Søraker 2009; Søraker 2012). Computers are usually not selected for one particular purpose (although in exceptional cases that may happen), but for many purposes, including web browsing, text processing, storing documents and other data, making spreadsheets, making PowerPoints, et cetera. Numerous cognitive proper functions may therefore coexist in one artifact. Furthermore, as computers may also be used for non-cognitive purposes such as playing music or online shopping, cognitive and non-cognitive functions can coexist in one artifact (see also Brey 2005).

Other cognitive artifacts do not have a history of cultural selection, but obtain their cognitive function through improvised uses. Sometimes such improvised uses are one-offs, e.g., when I am in a cafe and suddenly have an important idea which I quickly write down on a napkin before I forget it. Napkins have not been selected by their users to store information, but to help them clean or absorb liquid. However, during improvisation we can offload information onto the napkin (or any other artifact that affords writing on it[19]), thereby attributing a *cognitive system function* to the napkin. More specifically, as this is a one-off, it may be argued that the napkin becomes a cognitive artifact with an *idiosyncratic cognitive system function*. We may also attribute idiosyncratic cognitive system functions to a set of artifacts. Consider Norman's (1993) example of trying to explain how an accident happened by using everyday artifacts such as pencils and a paperclip as stand-ins for the objects (cars, dog) they represent. Pencils and paperclips have neither been designed nor selected to function as stand-ins for other objects, but during improvisation we may attribute cognitive functions to a set of initially non-cognitive artifacts.

In some cases, we improvise cognitive artifacts with more consolidated system functions, i.e., system functions that are more entrenched than mere one-offs[20]. Beach's (1988) study of bartenders who structure distinctively shaped drink glasses such that

[19] Although it does not need to be an artifact. The human hand is sometimes used to write on, e.g., a date of a deadline or a phone number. Jointly, the hand and the symbols written on it form a human-made information structure.

[20] In chapter 5, the degree of entrenchment is presented as one of the dimensions that are important to conceptualize the integration between embodied agents and cognitive artifacts.

they correspond to the order of the drinks is a good example. Due to this particular use of drink glasses, the bartenders do not have to remember the order of the drinks, but offload it onto their work-environment. Drink glasses are intended by their designers and selected by their users to contain their drink, not as mnemonic aids. However, this particular mnemonic use of drink glasses is (relatively) widespread practice for bar tenders, which are a small cultural group. This mnemonic use of drink glasses may therefore be seen as an *idiosyncratic ongoing cognitive system function*.

In other cases, we improvise cognitive artifacts with even more consolidated system functions. Most people put everyday artifacts in unusual locations such that they function as reminders. Leaving a rented DVD on your desk (or some other place) as a reminder to bring it back to the video store is a case in point. Such improvised mnemonic uses of everyday artifacts are quite common, e.g., leaving an empty milk bottle on the kitchen dresser as a reminder to buy milk, putting an article one has to read on top of the pile on one's desk, or tying a string around one's finger as a reminder for some action or event. In these cases, the location of the artifact is deliberately unusual such that it functions as external memory. In other cases, artifacts are put in a location that is deliberately usual. Some people always put their car keys on their hall table (or some other location) such that it is part of their behavioural routines, ensuring that they do not forget where they have put their car keys[21]. Such improvised uses of artifacts are widespread cultural practice and may therefore be referred to as *standardized ongoing cognitive system functions*.

The above examples show that we opportunistically use artifactual objects and structures for cognitive purposes, thereby improvising a variety of system cognitive artifacts, ranging from those with idiosyncratic exaptations (e.g. writing an idea on a napkin) to those with standardized ongoing exaptations (e.g. consistently leaving one's car keys on a certain spot). Due to these improvised uses of initially non-cognitive artifacts, we should conceive of the functional kind of cognitive artifacts as more inclusive than merely proper cognitive artifacts. In this sense, I agree with Hutchins that we should not only focus on "*designed* external tools for thinking" (Hutchins 1995, p. 172, italics added), not only because it encourages us to overlook cognitive techniques and external structures that are not designed, but also because it encourages

[21] This arguably happens with many artifacts.

us to overlook improvised tools for thinking. Thus, in order to increase explanatory scope and to better understand and classify a larger set of situated cognitive systems, we need to look at both proper and system cognitive artifacts.

2.3 Multiple Usability, Multiple Realizability, and Malfunction

Cognitive system functions vary between one-offs for individuals and widespread uses for a community of users, and everything in between. Furthermore, some of these system functions are ongoing additions to the proper function(s) of the artifact. Drink glasses, DVDs, hammers, and car keys are still used for their proper function, whereas napkins, empty milk bottles, and strings are usually thrown away after they have fulfilled their system function.

Conversely, proper cognitive artifacts can also be used for other functions. I may, for example, put some money between the pages of a textbook in my bookcase so that no one can find it, I may use a rolled-up newspaper as a flyswatter, a ruler to homogenize paint in a newly opened can, an abacus as a kids' toy, or some books to support my computer screen. Therefore, like most artifacts, (proper) cognitive artifacts are *multiply usable*, because their physical structure affords more than one particular use. So we can use a single (proper) cognitive artifact for a variety of functions, but we can also use a variety of (proper) cognitive artifacts for a single function. If, for example, I need to perform a difficult calculation, I may use pen, paper, and external numerals, an abacus, a pocket calculator, a spreadsheet program, a slide ruler, or even Charles Babbage's difference engine. These artifacts have rather different physical and informational structures, but on a course-grained (or macro) functional level of abstraction, they have the same function. If we were to ask: what are these devices for? Then the answer would most likely be: to perform calculations or to help us perform calculations. So it may be claimed that their macro-function is to (help us) calculate. Hence, like most artifact functions, cognitive functions are, at least on a macro functional level of abstraction, *multiply realizable,* i.e., different physical and informational structures can be used for the same cognitive function.

However, on a fine-grained (or micro) functional level of abstraction, there are differences in how these devices perform their function. When using pen, paper, and numerals, for example, most of the computation or information-processing is performed by an embodied brain. There may be different kinds of information-

processing at work when doing calculations with pen and paper and it may not always be easy to measure the amount of computation. When breaking down difficult calculations into easier ones such as, e.g., 3x7 or 5x5, some people may just remember the outcome rather than to actually compute it, while others may perform the easier calculations in their head. Either way, information is processed (remembered or calculated) by a human agent and not by an artifact. The external numerals function as to aid or supplement working memory and to structure the task space by decomposing the cognitive task in smaller and easier to perform parts. When using an abacus or slide ruler, the computation is done by an agent-plus-artifact system, i.e., manipulating the beads or slides *is* computation. The artifact's function, then, is to facilitate an analog computation performed by an agent-plus-artifact system. Finally, when using a pocket calculator or spreadsheet program, an agent merely provides the artifact with input and the computation is performed by the artifact. On a micro-level, the artifact's function is not to aid working memory, structure the task space, or to facilitate an analog computation, but to perform a digital computation.

In these scenarios, the artifactual elements have different micro-functions and the computation is distributed quite differently in each situated cognitive system. In the first scenario, there is mainly internal computation and the artifactual element is merely a medium for information storage; in the second scenario, there is agent-driven analog computation performed by a joint agent-artifact system; and in the third scenario, there is mainly digital computation performed by a computer. Moreover, using these different artifacts also requires a different set of cognitive techniques (i.e. interactive skills) and, consequently, the overall functional organisation of the situated cognitive system may be quite different. But, although different artifacts may have different micro-functions, may require different techniques, and result in different overall functional architectures of agents and artifacts, ultimately, the situated cognitive system has as its function to perform calculations (see also Clark 2008a; and see Gallagher & Crisafi (2009) on socially distributed systems exhibiting different composition but with similar functions regarding legal decision-making). One can therefore conceive of such systems as (to varying degrees) single integrated functional units. Thus when analysing the distribution and integration of cognitive functions, we may take a micro, macro or systems perspective. However, a proper analysis of situated cognitive systems should take all these perspectives into account.

The phenomena of multiple usability and multiple realizability of function are important properties of physical structures and are, in a sense, two sides of the same coin (Preston 2009). Certain physical structures can be used for multiple purposes and multiple physical structures can be used for the same purpose, although this does depend on one's functional level of abstraction. Another property of physical structures is that they can break down. When that happens, the artifact is said to *malfunction*. So, strictly speaking, malfunction is not something going wrong with the cognitive function itself, but with the physical structure of the artifact. Cognitive artifacts can also misrepresent their content, which is a form of malfunction, too, and is discussed in the next chapter. Malfunction can thus happen when physical structures breaks down or when representational structures misrepresent their content. When it concerns the breakdown of physical structures, it may be due to misuse, manufacturing defects, poor design, accident, or simply wearing out as a result of normal usage. To prevent this from happening, artifacts need maintenance, repair, or in some cases need to be rebuild. In extreme cases, when repair is not worth the effort, a cognitive artifact may be recycled, which happens with, e.g., computers and smartphones.

3. Classification Systems in Archaeology

Having identified the target domain and described some functional properties of physical structures, I now look at how artifacts are classified in archaeology. Philosophy of technology has not focussed much on developing classification systems for artifacts. Archaeology, by contrast, has long-standing and robust methods for classifying their objects of interest. For this reason, I borrow from this fields, in order to develop a taxonomy of cognitive artifacts, which is done in the next chapter.

3.2 Artifact Classifications in Archaeology

There is no classification system for artifacts in philosophy of technology. One possible reason for this is that there is no unified science of artifacts, nor is there a unified philosophical theory of artifacts (see also Dipert 1993), although some useful steps have been made by Houkes & Vermaas (2010), Kroes (2012), Preston (2013), and Vries, Hansson, and Meijers (2013). These theorists, however, do not focus on categorizations but on other metaphysical issues concerning artifacts such as function or normativity. Categorizations that do exist for the artifactual domain are usually restricted to a relatively small set of artifacts. In archaeology, for example, artifacts are often

categorized on the basis of their morphological, historical, or functional properties. Morphological classifications contain categories that are defined by observable physical properties such as size, shape, colour, and decoration. An archaeologist may, for example, classify pottery in terms of decoration, spear points in terms of size, stone tools in terms of shape, and so forth. Such classifications are developed mainly for purposes of description and comparison, for example to map out the variation of artifacts unearthed at a particular archaeological site, or to compare different classifications of artifacts from different sites.

Historical classifications categorize artifacts in terms of the historical period in which they have been manufactured. Thus a hand-axe produced during the Stone Age is categorized as a Stone Age tool, a bronze spear point produced in the Bronze Age is categorized as Bronze Age tool, and so forth. In these cases, artifacts are categorized based on their historical properties, not on their observable physical properties. However, in practice, historical classifications often use stylistic attributes, as these are most likely to vary from time to time. As there are many kinds of tools manufactured during certain historical periods, this way of categorizing results in rather broad categories without any morphological or functional specification. In their excellent book, anthropologist William Adams and philosopher of science Ernest Adams write: "all of the different things found by the archaeologist - tools, containers, houses, pictographs, sites, and whole "cultures" - may at times be classified historically" (Adams & Adams 1991, p. 220). Historical classifications are often developed to provide insight and clarity to how artifacts have changed and developed over time, for example by comparing different historical classifications.

Functional classifications, finally, contain categories defined on the basis of the presumed purpose for which the artifact was used, i.e., its function. However, as Adams & Adams aptly point out:

"function is nearly always an inference by the archaeologist, based on observable morphological (intrinsic) features. It would therefore be more nearly correct to say that a functional type is one defined by a limited set of intrinsic variables - specifically those that give evidence of the use or intended use for which the artifacts were made" (Adams & Adams 1991, p. 222).

Function is thus inferred from a limited set of observable physical properties, which is rather speculative because

"pottery and many other kinds of artifacts usually exhibit attributes that cannot be related to specific functions, but that are too conspicuous and too variable to be ignored... the making of purely functional artifact typologies is really very rare, at least in prehistoric archaeology. The usual practice is, rather, to attempt a functional interpretation, *a posteriorly* to types that have been defined phenetically" (Adams & Adams 1991, p. 222).

Consequently, functional classifications of artifacts unearthed at archaeological sites are often rather general and contain categories such as "agricultural tools", "household implements", "weapons", "clothes", and so forth. The main purpose for which functional classifications are developed is to help reconstruct the activity patterns and cultural practices of the humans that inhabited a certain area. For when the function of an artifact is known, one also knows what its user(s) did with it. Different classifications serve different purposes: e.g., mapping variation (morphological), understanding how artifacts have developed over time (historical), reconstructing cultural practices (functional), or a combination of these[22].

A distinction can be made between two kinds of classifications: typologies and taxonomies (Adams & Adams 1991). This distinction concerns the complexity and the amount of levels of abstraction in the classification. The categories (or types) in a typology are considered to be more or less at the same level of abstraction, whereas taxonomies have categories (or taxa) on more than one level of abstraction. In other words, taxonomies exhibit more complexity and are hierarchical, multi-level classification systems, whereas typologies are typically less complex, one-level classification systems[23]. Depending on each specific case, morphological, historical, and functional classifications may either be typological or taxonomic. A final distinction I

[22] For overviews on the typology debate in (philosophy of) archaeology, see Adams & Adams (1991, chapters 22 & 23), Wylie (2002, chapter 2) and Read (2007).

[23] In terms of this distinction, the current categorizations of cognitive artifacts that have been discussed in the previous chapter (Donald, Norman, Brey, Nersessian, and Sterelny) are typological, rather than taxonomic, as they have one level of abstraction, i.e., they do not make further distinction within existing categories.

want to borrow from Adams & Adams (1991) is one between basic and instrumental purposes. A classification serves a basic purpose when it is developed to better understand the material being classified and it serves an instrumental purpose when it is developed to better understand something other than the classified material. So, for example, a morphological classification may be seen as serving a basic purpose (as its goal it to map variation in artifacts), whereas a functional classification as serving an instrumental purpose (as its goal is to reconstruct cultural practices).

4. Concluding Summary

This chapter began with identifying the target domain by defining the category of cognitive artifacts as a functional kind. More specifically, this functional kind includes cognitive artifacts with both proper and system functions. Those with proper functions have a history of cultural selection, whereas those with system functions are improvised uses of initially non-cognitive artifacts. Next, different classification systems in archaeology were outlined: morphological, historical, and functional classifications. I have now demarcated the target domain and provided some of the tools to further classify and taxonomize this target domain, which is done in the next chapter by drawing on artifact categorization in archaeology.

4

A Taxonomy of Cognitive Artifacts

1. Introduction

In the previous chapter, I identified the target domain and looked at how and why objects may be categorized. In this chapter, I propose and develop a method to categorize the artifactual element in situated cognitive systems, drawing on artifact categorization in archaeology. In chapter 2, I demonstrated that current categorizations focus on representational artifacts. They largely neglect non-representational or ecological artifacts and also tend to focus on proper cognitive artifacts and often (though not always) overlook system cognitive artifacts. Given that the functional kind of cognitive artifacts includes more than just proper representational artifacts, current categorizations therefore have a significantly smaller and less inclusive target domain. Moreover, current categorizations are cognition-centered, i.e., they start with human cognition and then work their way towards the (proper) representational artifacts that

agents deploy to aid their cognition and realize their cognitive purposes. As a result, most of these approaches (with Norman and Donald as notable exceptions) do not have a great deal to say about the particular informational properties of artifacts deployed in performing cognitive tasks.

In this chapter, I propose an alternative approach to categorizing cognitive artifacts in which an information-centered approach is taken, i.e., I take as my point of departure the specific properties of cognitive artifacts and then taxonomize them on the basis of those properties. This focus on informational properties of the artifact may raise two worries: (1) that the larger situated cognitive system is not important and (2) that those properties are intrinsic to the artifact. Regarding the first worry, situated cognitive systems have components that interact, transform each other, and are to varying degrees integrated into a larger cognitive system. One part of a possible strategy to better understand the larger system is by decomposing it into its components and by conceptualizing the particular cognitive properties of those components. This strategy is helpful because those properties largely determine how the situated system performs its cognitive tasks and to what degree the components are integrated and transform each other, which is the topic of part II of this thesis.

To give a brief preliminary example, one species of cognitive artifacts that are identified in the taxonomy are indices (e.g. a thermometer). Indices are causally connected to their target system. The information an index provides cannot be changed by the user and so the information flow between user and artifact can only be one-way, which limits the degree of integration between user and artifact. The informational content of other species of cognitive artifacts such as icons and symbols, in contrast, can be altered by the user and therefore allows two-way or reciprocal information flow. This allows the artifact to be integrated much deeper with the cognitive system of its user. Better understanding the informational properties of the artifact is thus helpful for better understanding the degree of integration between agent and artifact. Furthermore, it is important to point out that cognitive artifacts are embedded in a larger practice. Ultimately, the goal is to better understand the larger practice, which is achieved by better understanding the informational properties of artifacts and the degree of integration.

Therefore, I first look at the specific cognition-aiding properties of artifacts, i.e., their informational properties (this chapter) and then develop a multidimensional framework for conceptualizing how those informational properties and their users are integrated into systems that perform cognitive tasks (chapter 5). Next, I apply this multidimensional framework to cases of agent-artifact integration in molecular biology laboratories (chapter 6). I thus try to understand the larger situated cognitive system by first categorizing artifacts and then by conceptualizing how these artifacts and their users are integrated into a larger system. The larger situated cognitive system is thus important in the overall goal of this thesis. Regarding the second worry, as I argued in the previous chapters, cognitive artifacts are neither defined by intrinsic properties of the artifact nor by the intentions of the designer, but by their function, which is established by the intentions of the user and by how it is used. Informational properties are thus not intrinsic to the artifact, but are sometimes attributed to them by their users.

I proceed as follows. I start with developing a method to taxonomize the artifactual element in situated cognitive systems by drawing on artifact categorization in archaeology (section 2). In this taxonomy, I distinguish between three levels or taxa, those of family, genus, and species. The family is an overarching category that includes all cognitive artifacts without further specifying their informational or functional properties. I then distinguish between two genera: representational and ecological cognitive artifacts. These two genera are further divided into species. In case of representational cognitive artifacts, three species are identified, those that are iconic, indexical, and symbolic. In case of ecological cognitive artifacts, two species are identified, those that are spatial and dynamic (section 3). Finally, a concluding summary is given (section 4).

2. Methodology

Which properties of cognitive artifacts are relevant for developing a taxonomy? In archaeology, artifacts are categorized on the basis of their morphological, historical, or functional properties. Depending on the goal for which the artifacts are categorized, e.g., mapping variation, understanding how artifacts have developed over time, or reconstructing cultural practices, one may prioritize morphological, historical, or functional properties. The goals of developing a taxonomy of cognitive artifacts are

twofold: (1) to provide structure and systematicity in the variety of cognitive artifacts, which should result into (2) a better understanding of the epistemic interactions and integration between embodied agents and different kinds of cognitive artifacts. In other words, the goals here are to map current variation and to understand the cultural practices regarding the use of cognitive artifacts. So, in terms of Adams & Adams (1991) distinction between basic and instrumental purposes, the proposed taxonomy serves both a basic and an instrumental purpose, because it is developed to better understand the material being classified and to better understand something other than the material being classified, namely, situated cognitive systems. With this basic and instrumental purpose in mind, we should prioritize morphological and functional properties, rather than historical properties[24].

It is essential to further clarify the relation between the artifact's morphology and its function. A cognitive function supervenes on a particular kind of morphological properties, namely, informational properties. In other words, cognitive functions are established only when an artifact exhibits information that is used in performing some cognitive task. If there is no information, then there are no cognitive functions. Exhibiting task-relevant information is thus a necessary and sufficient condition for being a cognitive artifact[25]. Hence, in order to map the variety of cognitive artifacts, the relevant properties for developing a taxonomy are their informational properties, because these really matter to the user in performing a cognitive task and are necessary for establishing their cognitive function. In the previous chapter, a distinction was made between representational and ecological artifacts. This distinction between representational and ecological artifacts will be the point of departure for developing the taxonomy and is further developed below.

---

[24] However, given that proper functions of cognitive artifacts are established through a history of cultural selection, historical properties are sometimes implicitly reflected in the taxonomy, but they are not prioritized in functional kind membership, because a history of selection is not necessary for attributing cognitive functions to artifacts. Furthermore, I want to point out that developing an historical taxonomy of cognitive artifacts, i.e., one that maps their historical development, would be an interesting and important research project and Donald's (2010) attempt to list generic exogram systems in roughly chronological order would be a useful first step (see table 1 in chapter 2).

[25] Even when the artifact exhibits task-relevant information, they may still malfunction.

The proposed taxonomy is hierarchical and starts at a very general and inclusive level (i.e. the level of a functional kind) and gets increasingly more specific and less inclusive when one goes deeper into the taxonomy. I distinguish between three levels of abstraction or taxa, those of family (i.e. functional kind), genus, and species (see figure 1 below). The family includes all cognitive artifacts (i.e. proper, system, representational, and ecological) without further specifying their informational properties. On the second level in the taxonomy, i.e., the taxon of genus, I distinguish between two genera: representational and ecological artifacts. On the third level, these two genera are further divided into species. In case of representational artifacts, three species are identified, those that are iconic, indexical, and symbolic. In case of ecological artifacts, two species are identified, those that are spatial and dynamic. Finally, particular instances of artifacts are identified, i.e., token cognitive artifacts.

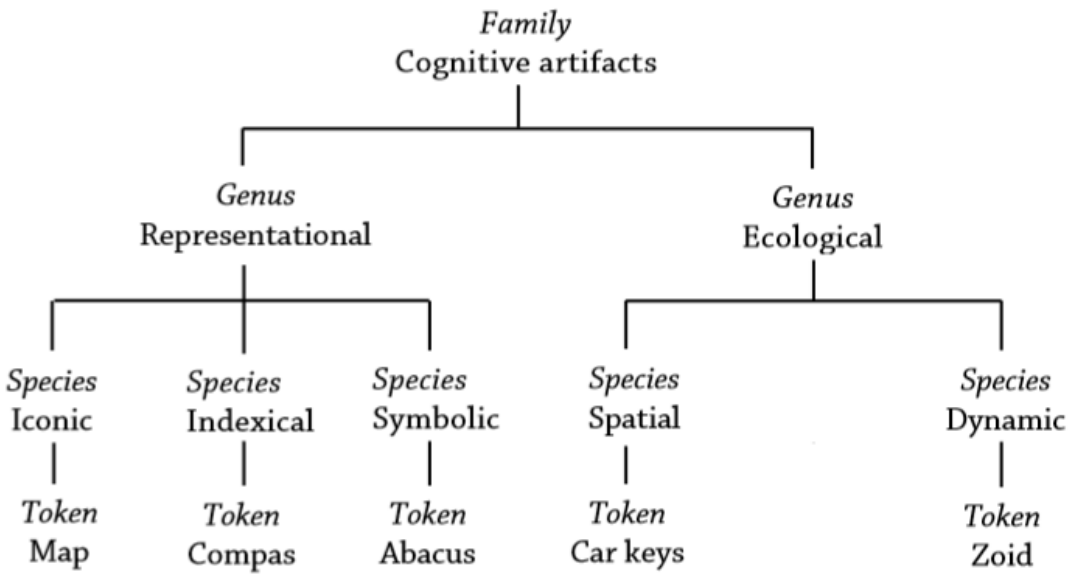

Figure 1. A diagram of a taxonomy of cognitive artifacts.

The logical structure between the categories in the taxonomy can be described as follows. The relation between family and genus is that the latter is necessarily part of the former, but not the other way around. Hence, a representational cognitive artifact is necessarily a cognitive artifact, but not all cognitive artifacts are representational. Families or functional kinds are high-level ontological categories with two genera as members. A genus is a mid-level ontological category and artifacts in it are defined by exhibiting distinct informational properties, which are representational or ecological, whereas a family is purely defined by its function without further specifying informational properties. The two genera are not mutually exclusive, as there are cases

in which representational and ecological information jointly contribute to performing a cognitive task. Genera have species and, like genera, species are defined by their distinct informational properties. The three species of the representational genera are characterized by being iconic, indexical, or symbolic and are not mutually exclusive, but one of those properties is typically predominant. The two species of the ecological genera are characterized by being spatial or dynamic.

Taxon	Map
Family	Cognitive artifact
Genus	Representational
Species	Iconic
Type	Map

Table 1. Taxonomic ranking of a map.

In the above taxonomy, family, genus, and species membership are monothetic, which means that one property is the determining factor for membership. In case of family, having a cognitive function is the determining property for membership. In case of genus and species, exhibiting a particular informational property is the determining factor for membership. So this taxonomy combines functional and informational properties for category membership. Finally, upward deduction in the taxonomy is valid, but downward deduction is not. So, when membership at the type or species level is known, one can deduce the membership of the higher levels, but not vice versa.

3. Representational Cognitive Artifacts

I now further conceptualize the genus of representational cognitive artifacts and distinguish between three species, those that are iconic, indexical, or symbolic. We have seen that some artifacts have representational properties. But what exactly are representational properties? To answer this question, we need to take a closer look at what a representation is. A useful starting point is the work of American pragmatist Charles Saunders Peirce. Peirce argued that a representational system involves a *triadic* relation between an *interpretant* (a cognitive effect in a human agent, usually understood in terms of the agent thinking about the object in question), *sign*

(representation or representational vehicle[26]), and *object* (content or represented world). There is a potential terminological issue here. Peirce refers to the content of a representation as its object. But sometimes that object is literally an existing physical object, e.g., a tree, house, or chair. To avoid this issue, I use Cummins' (1996) terminology of "target" and "content".

Signs or representations thus represent something. John Haugeland has aptly put it as follows: "That which stands in for something else in this way is a *representation*; that which it stands in for is its *content*; and its standing in for that content is *representing* it" (Haugeland 1991, p. 62, original italics). A defining property of a representation is thus that it stands in for something else, i.e., it has aboutness or representational content. Consequently, a representation is an information-structure that provides us with information about something external to the representational vehicle itself[27]. Representational vehicles can represent different kinds of information: an object, property, category, relationship, state of affairs, or a combination of these. For instance, a blueprint of the Singer Building in Manhattan represents a particular object. A red sticker on a can of paint represents a property of the paint inside the can. The word "buildings" represents a category of objects. Pythagoras' theorem, $a^2 + b^2 = c^2$, represents a relationship in Euclidian geometry between the three sides of a right triangle. And a picture of a cityscape, say, Manhattan, represents a state of affairs. These examples show that there are different forms of representation, ranging from blueprints to words and from formulas to pictures. Importantly, not all properties of the representational vehicle are relevant for its content. The red sticker on a can of paint, for example, may be square and made of plastic. But these properties have nothing to do with its content, for which only its colour is relevant.

3.1 Icons, Indices, and Symbols

According to Peirce, a representational system is irreducibly triadic, but for analytical purposes, we can decompose Peirce's triadic relation into two dyadic relations; one between the representation and its target, and one between a human agent and the

[26] The terms "representation", "representational vehicle" and "vehicle" are used interchangeably throughout this dissertation.

[27] Although some representations provide information about themselves. For example, Rene Magritte's painting, *The Treachery of Images,* contains a representation of a pipe and the sentence "Ceci n'est pas une pipe", which is French for "This is not a pipe." Some representations are thus self-referential.

representation. Let us first have a look at the representation-target relation. Peirce (1935a, 1935b) distinguished between three possible grounds for representation, which are iconic, indexical, and symbolic (see also Burks 1959; von Eckhardt 1995; Atkin 2010). An icon is relevantly similar or isomorphic to what it represents. Straightforward examples are maps, radar systems, scale models, and blueprints, which are iconic because there is a high structural isomorphism between the content of the map, radar system, scale model, or blueprint and their targets[28]. These artifacts have been intended by their designers and selected by their users to function as icons. They may therefore be seen as proper icons. There are, however, also improvised or system icons. Consider (again) Norman's (1993) example of trying to explain how an accident happened with the aid of pencils and a paperclip as stand-ins for the objects (cars, dogs) they represent. Pencils and paperclips may not in any way resemble cars and dogs, but the order and location in which they are placed is isomorphic to the order of the cars and dog in the actual accident. So only some elements in this system icon are isomorphic to their targets.

Note that isomorphism is a quite general concept and because of this generality it can be pushed rather far, as to include representations such as graphs and other diagrams (see also Haugeland 1991). Consider a line graph representing the amount of carbon dioxide in the earth's atmosphere plotted against the time. This graph represents the relationship between carbon dioxide and time. The amount of carbon dioxide has been increasing over the last hundred and fifty years or so and thus the line graph will show a line going up. There is indeed some kind of isomorphism between an increasing amount of carbon dioxide particles in the atmosphere over a certain period of time and a line going up, but this isomorphism is of a very different kind than, for example, the isomorphism between a map and its target which is much stronger. Thus, in some cases, we can no longer speak about resemblance between target and icon and so isomorphism between icons and their targets may be abstract.

Indices have a direct causal connection between the index and its target. There is, for instance, a direct causal connection between the direction of the wind and the direction of a weathervane. If the direction of the wind changes, then the direction of the weathervane automatically changes, too. So the position of the weathervane is an index

[28] Icons may also be natural or biological, e.g., the amount of tree rings resembles the age of the tree.

for the direction of the wind. Note that it is also partly iconic, since there is an isomorphism between the direction of the wind and the direction of the weathervane. Other examples of indices are thermometers, compasses, scales, voltmeters, speed dials, barometers, and many other measuring instruments. These are indices because there is a direct causal connection between the representational state of the index and its target. In these indices, the following things are causally connected: temperature and the expansion of mercury in a thermometer, the location of the North Pole and the direction of an arrow in a compass, the mass of a particular object and the reading on a scale, the amount of volts in an electrical current and the reading on a voltmeter, the speed of a car and the reading on a cars' speed dial, and atmospheric pressure and the reading on a barometer. These are all proper indices, as they are intended by their designers and selected by their users to function as indices[29]. System indices also exist but are much less common than, e.g., system icons. This is so because most indices have a rather complex physical structure which is hard to improvise, although there are cases of improvised thermometers which can be made of a plastic bottle, some clay, a straw, coloured water, and a piece of paper.

The last kind of representations that Peirce identified are symbols[30], which acquire their meaning and content through shared use, agreement, and logical rules. Typical cases are words and sentences in natural language, numerals, or symbols in mathematical and scientific formulas. The form or structure of symbols is quite often (though certainly not always) arbitrary. There is, for example, nothing intrinsic in the structure of the word "buildings" that makes it represent the category of buildings, there is likewise nothing intrinsic in the structure of the numeral "4" that makes it represent 4 units[31],

[29] I would like to briefly point out an example of a malfunctioning index that nevertheless has a history of cultural selection, but cannot actually perform its function. Some funfairs and arcades have machines that supposedly measure one's sex appeal, love abilities or romantic feelings for someone. Such machines are sometimes called Love Tester Machines. They work by measuring the moisture and/or temperature of one's hand and based on that give an induction of one's sex appeal, love abilities or romantic feelings for someone. These machines obviously do not measure these things, but they are nevertheless re-produced (an old love tester can be found in the Musée Mécanique in San Francisco).

[30] There is a potential terminological issue with the notion of symbol here. Other theorists such as Donald (1991, 2010) use the notion of symbol (e.g., in the idea of an external symbol system) to denote all representational systems including what Pierce calls icons and indices. So an external symbol system is much broader than Peirce's notion of symbol in that it includes iconic and indexical representations.

[31] Although this is different in the Roman numeral system, see Zhang & Norman (1995).

and there is nothing intrinsic in the structure of the sign for wavelength "λ" that makes it represent wavelength. There is no resemblance or causal connection between symbols and their target. It is shared use, agreement, and logical rules that establishes the meaning of symbols, not their representational structure. Sometimes other kinds of representations have symbolic properties as well. Certain icons such as diagrams cannot function as icons merely by their isomorphic relations to their target. There needs to be an agreement or convention that indicates which elements of the icon are to be interpreted as being isomorphic to its target. Measurement outcomes of certain indices are also partly symbolic, e.g., temperature may be measured in degrees Celsius or Fahrenheit, depending on one's geographical location and convention. Thus, as Peirce recognises, any particular representational vehicle may display a combination of iconic, indexical, and symbolic characteristics. Consequently, a useful way of conceptualizing Peirce's trichotomy is by seeing a token representation as *predominantly* iconic, indexical, or symbolic (Atkin 2008).

Like icons and indices, symbols function so as to represent their target. Most current symbols have a fairly long history of cultural selection and may therefore be seen as proper symbols. Because the meaning of symbols is defined in virtue of shared use, agreement, and logical rules, it is possible to develop improvised or system symbols that may over time become consolidated into proper symbols. We may image a scientist, engineer, designer, or some other individual(s) inventing new symbols, symbolic structures, or symbolic systems, e.g., a symbol for a newly discovered phenomena, a new computer programming language, or a new musical notation system. Initially, these new symbols only have system functions, as they lack a history of cultural selection, but depending on how often they are used and how successful they are in spreading to other cultural groups, they may become consolidated into proper symbols. This is arguably what happened to most current proper symbols and to proper icons and indices as well. Preston argued that the notion of system function is important because it explains the development from the simple exaptive use of naturally occurring objects as hammers, digging sticks, and so forth, to the high-tech world in which we now live. It likewise explains the development of representational systems that at some point must have started as improvised representational structures.

Finally, there are also system cognitive artifacts containing proper symbols. Offloading information in the format of written language onto a napkin is a case in point. The

written language itself is a proper symbolic structure, but when offloaded onto an artifact that is neither designed nor selected to function as an information storage device, the artifact together with the proper symbols offloaded onto it, become a system cognitive artifact.

3.2 Forms of Representation

One of the advantages of Peirce's trichotomy is that it allows us to taxonomize all forms of representation as predominantly iconic, indexical, or symbolic. Donald Peterson, in his edited book, *Forms of Representation: An Interdisciplinary Theme for Cognitive Science*, correctly points out that we use a great variety of representational systems, including:

"algebras, alphabets, animations, architectural drawings, choreographic notations, computer interfaces, computer programming languages, computer models and simulations, diagrams, flow charts, graphs, ideograms, knitting patterns, knowledge representation formalisms, logical formalisms, maps, mathematical formalisms, mechanical models, musical notations, numeral systems, phonetic scripts, punctuation systems, tables, and so on" (Peterson 1996, p. 7).

Peirce's trichotomy can be used to categorize these representational systems. Algebras, alphabets, computer programming languages - knowledge representation, logical, or mathematical formalisms – musical notation systems, numeral systems, phonetic scripts, punctuation systems, and tables are predominantly symbolic, as they mainly acquire their meaning through logical rules and agreement. But some of these representational systems also display isomorphism with their targets. For example, in the case of musical notations, the order of the symbols for the notes is isomorphic with the order of the notes in the actual piece of music. Likewise, certain numeral systems such as, e.g., the Roman, Babylonian, and Chinese rod systems have isomorphic elements, as some of their numerals are isomorphic to the amount of units they represent. Three units, for example, are represented as three vertical stripes in those numeral systems.

Animations, architectural drawings, choreographic notations, computer models and simulations, diagrams, flow charts, graphs, ideograms, knitting patterns, maps, and mechanical models are predominantly iconic, as they acquire their meaning through exhibiting some kind of isomorphism with their targets. However, some of these

systems, for example, architectural drawings, computer models, and diagrams may also contain symbolic representations such as words, sentences, and numerals. And, more importantly, there are sometimes rules for interpreting certain elements within those representations as isomorphic to their targets. For example, the legend on a map often contains rules and guidelines for interpreting certain elements in the map as being isomorphic to certain objects and structures in the world. Note that Peterson's list does not contain any indices. Finally, computer interfaces do not really belong in that list, as they are not a form of representation, but more a medium in which a variety of representational systems can be expressed, computed, and manipulated.

Thus, icons, indices, and symbols are species of the genus of representational cognitive artifacts and can account for all forms of representation. Because these artifacts are predominantly iconic, indexical, or symbolic, species membership is monothetic, i.e., one informational property determines species membership. However, there may be cases in which it is hard to determine which informational property is predominant. In certain pictographic languages such as Egyptian hieroglyphs, particular characters have a strong isomorphism to their target, other characters have a less obvious isomorphism to their target, and yet other characters have no isomorphism to their target. A sentence, or series of semantically connected hieroglyphs, may combine isomorphic (icons) and non-isomorphic (symbols) hieroglyphs in such a way that it hard to determine whether the overall representational structure is predominantly iconic or symbolic. Also, because isomorphism between an icon and its target can be abstract, it may not always be possible to demarcate between an abstract isomorphic icon and a symbol.

3.3 Mind-Dependent Targets and Normativity

Representations do not necessarily have to represent things in the world that actually exist. Consider an architect's blueprint of a not yet existing structure. It makes sense to say that the blueprint is a genuine iconic representation, even before the structure has been build. One could say that the blueprint is a representation of a possible world, where the likelihood of that possible world coming into existence is significant. Similarly, but conversely, an architectural blueprint of the Singer Building, a demolished landmark in Manhattan, is still an iconic representation of an iconic building, even if it no longer exists. However, the targets of the representational vehicle need not ever have existed or have a high likelihood of coming into existence. The word

"unicorns" represents the category of unicorns. But unicorns do not have the same ontological status as real biological organisms have. They never have, and probably never will. They are, however, existing entities in European folklore and mythology and so the word "unicorns", or an icon of a unicorn, has content. An architect's blueprint of a not yet existing structure or a no longer existing structure (an icon) and the word "unicorns" (a symbol) have representational content, but their content is not about a currently existing state of affairs. Therefore, the content of a vehicle may represent things that have existed, will exist, or only exist in the human mind.

Finally, representations can be descriptive or prescriptive, i.e., they can describe existing or non-existing entities, but they can also prescribe what actions we should undertake. Consider icons such as no-parking signs that tells us not to park at a certain spot, or no-smoking signs that tell us not to smoke. Such representations do not represent existing things in the world, but prescribe how to act. Their content is therefore normative in the sense that the representation depicts a behavioural norm to which an agent has to conform. Note that this is only true for those that would have otherwise parked or smoked. Written language is also often used to convey prescriptive content, including documents containing professional codes, law books, and elaborate normative theories in ethics such as utilitarianism or deontology. Finally, it seems that only icons and symbols can be used to convey prescriptive content. It is difficult for an index, which has a direct causal connection to its target, to explicitly represent prescriptive content. It may perhaps implicitly represent prescriptive content, e.g. an index that directly measures the amount of carbon dioxide in the earth's atmosphere may be interpreted as a warning sign to undertake action, i.e., to reduce our carbon emissions.

3.4 Structure-Function Relations

In connection to structure-function relations, Peter Kroes and Anthonie Meijers point out that "Technical artefacts can be said to have a dual nature: they are (i) designed physical structures, which realize (ii) functions, which refer to human intentionality" (Kroes & Meijers 2006, p. 2, see also Kroes and Meijers 2002). Physical structures are often said to be mind-independent and can be described by the laws of physics, whereas artifact functions are mind-dependent (i.e. they require for their existence human intentionality) and thus require an intentional description. So, in order to properly describe artifacts, Kroes and Meijers argue, we need to somehow combine physical and

intentional descriptions. Pieter Vermaas and Wybo Houkes (2006) argue that the notion of artifact function is helpful here because it is a "conceptual drawbridge" between the physical and intentional realms. Functions tie the physical and intentional realms together, i.e., functions necessarily need both physical structures and intentional human agents that design, select, improvise, and interact with those physical structures. In this sense, functions can be seen as emergent properties of intentional agents interacting with human-made physical structures that have a particular effect. Drawing on this insight, a model of the emergence of cognitive functions can be sketched.

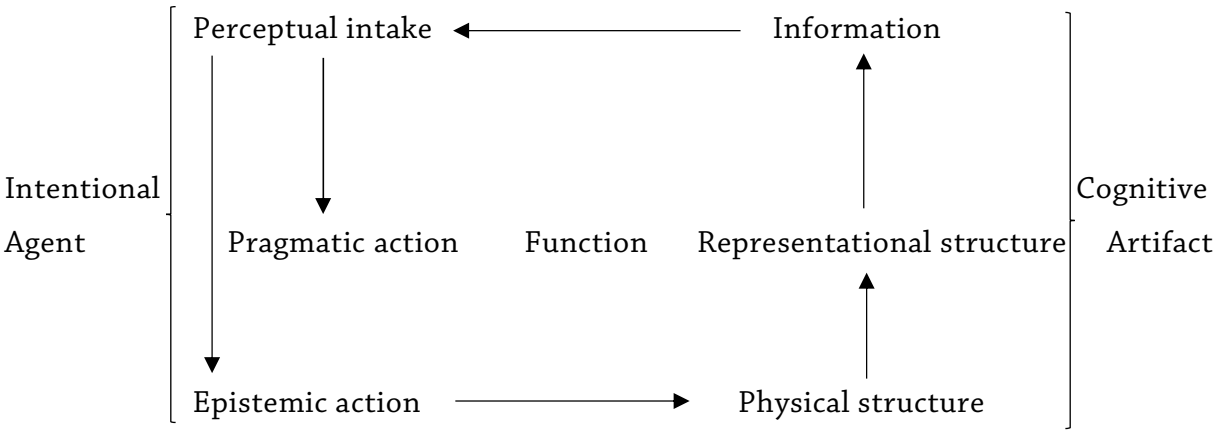

Figure 2. Model of interaction between intentional agents and representational cognitive artifacts.

The above figure presents a somewhat simplified model of cognitive function as an emergent property of the interaction between intentional, embodied agents and cognitive artifacts. To briefly illustrate this model, I use the example of navigating with a map. We interact bodily with the physical structure of the map, for example orienting it such that the information it contains becomes available to our perceptual systems. Such epistemic actions have as their goal to make available task-relevant information (Kirsh & Maglio 1994; Clark & Chalmers 1998; compare Loader 2012). The representational structure of the map contains a large amount of information, but only a small part of it is used to perform some cognitive task. A map of Sydney's central business district, for example, may contain an elaborate representation of all the streets, parks, landmarks, train stations, bus stops, and so on. But to navigate, only a relatively small part of all the available information is relevant. The information that is actually used is the task-relevant information. This information is perceived and then

processed by internal systems, either to guide a pragmatic action (e.g., walking towards the Harbour Bridge) or to guide further epistemic action (e.g., reorienting the map). This model shows how cognitive functions emerge out of intentional interactions with physical, information-bearing structures.

Because functions define artifacts and tie the physical and intentional realms together, it is important to better understand how they are established. One way to do this, is by taking a closer look at the following relation:

(1) physical structure → representational structure → information → cognitive function

We have already seen how information supervenes on representational structures in the previous subsection, i.e., information obtains its meaning by being isomorphic to its target (icons), by having a causal connection to its target (indices), or by logical rules and agreement (symbols). In this subsection, therefore, I focus mainly on the relation between physical and representational structure, i.e. the first arrow in (1).

There are at least three different kinds of relationships between physical and representational structures. First, representational structures can be carried by or added onto a physical structure. Icons and symbols are often carried by paper, whiteboards, screens, or some other material. A map of the London subway system, for example, is printed onto a piece of paper, symbols (e.g. words, numerals) in a notebook or textbook are written down or printed onto a piece of paper, lecture notes are written onto a whiteboard or projected onto a screen, and the metric scale and numerals are engraved into or painted onto a ruler. In such cases, physical materials such as paper, plastic, or some other material carry a representational information-structure. Other materials have (historically) also been used to carry external representations, including rock, papyrus, clay tablets[32], wax tablets, wood, animal skin, canvas, and even sand. What is important here is that the physical structure of the underlying material need to be such that it can sustain an external representation for a certain period of time. This depends on the cognitive purpose for which the artifact is deployed. Writing with one's

[32] The work of Lambros Malafouris (2004) is particularly helpful when understanding how information is stored in clay tablets.

finger in the sand on a beach may work for creating a simple drawing, but not for making a detailed enduring architectural blueprint which needs a fairly enduring material such as paper. Durability will be discussed in more detail in the next chapter. In the above mentioned examples, the relation between the physical and representational structure is one of carrying and, therefore, the structure of the physical material does not directly influence informational content, only the structure (i.e. form and shape) of the icons and symbols are relevant for informational content.

Second, representational structures can be directly constituted by physical structures. When this happens (part of) the physical structure of the artifact is identical to the representational structure it exhibits. There are two ways in which this can happen: statically or dynamically. Designers sometimes make scale models (i.e. icons) of the objects or structures they are designing, for example to test the aesthetic value, aerodynamics, or physical strength of their design. Scale models have a physical structure that is identical to their representational structure. There are usually no additional representations added onto a physical structure, as is the case with, e.g., maps, notebooks, textbooks, lecture notes, rulers, et cetera. However, in certain scale models, for instance those that are made to test the aesthetic value, there is colour added onto their physical structure, in which case there is an additional representational structure added to a physical structure. If that does not happen, then the physical form or structure of the artifact *literally is* its representational structure on which its information-structure supervenes. This is statically constituted because its physical and thus representational structure does not change. After the scale model has been made, its structure usually remains unaltered.

It is perhaps helpful to briefly point out that the relationship between representational structure and informational content is multiply realizable. Compare, for example, a detailed blueprint (either printed on paper or a digital version) and an accurate scale model of a token building. Let us assume that these contain exactly the same informational content about their target, i.e., size, form, ratios, colour, structural composition, and so on, presented in different representational formats. Although these icons may look different and from a phenomenological, user-centered perspective we may experience them differently, strictly speaking they do contain the same informational content. There are, however, differences in affordances of the different formats, which may have informational consequences. Similarly, Herbert Simon (1978)

argued that two different representational formats, for example propositional (symbolic) and diagrammatic (iconic), might be informationally equivalent, but not computationally equivalent. So they can contain the same informational content, but afford different (kinds of) computations.

In case of indices such as thermometers, barometers, and compasses, their physical structure and therefore also their informational content is dynamically coupled to their target. For example, the physical structure of a mercury thermometer is causally coupled to the temperature. In other words, mercury expands when temperature increases, which under normal conditions constitutes a linear relationship between temperature and the degree of expansion. In this case, the physical structure and state of the artifact (i.e. the diameter of the column and the particular expansion properties of mercury) is partly identical to its informational content, i.e., the temperature. Thermometers also contain a static temperature scale such that one can precisely see what temperature it is. Jointly, a static representational structure (i.e. the temperature scale which is carried by some material) and a dynamic physical structure (i.e. mercury in a column) constitute the informational content of the artifact. Similar relationships between physical structure and informational content are established in barometers and compasses, where expansion properties of an alloy partly determine the informational content of a barometer (i.e. atmospheric pressure) and magnetic properties of a compass arrow partly determines informational content of a compass (i.e. the cardinal directions). If the properties of the physical material were even slightly different, then the informational content would be different as well. Thus the target causally changes the physical structure and state of the artifact, which, together with a static representational structure, constitutes the informational content of the artifact.

Third, representational information can be computed or manipulated by a physical structure. This happens in both analogue and digital computational artifacts. Such artifacts have a physical structure that affords information-processing or computation. A slide ruler, for example, is a mechanical analog computer used mainly for multiplication and division, and to a lesser extend for calculating roots, logarithms, and trigonometric functions. Slide rulers contain a set of static logarithmic scales that can be manually manipulated such that a mark on the sliding strip is aligned with a mark on the fixed strip. The relative positions of other marks on the strips are then observed. Numbers aligned with the marks give the approximate answer to the calculation. By

manipulating the physical structure of the slide ruler, one automatically manipulates its representational structures and thus its informational content, in that way performing analog computations.

Analog computational artifacts are often designed such that they can perform one type of computation (e.g. mathematical computations) often consisting of one type of representations (e.g. numerals). Their physical structure severely limits how and what kind of information can be manipulated and processed. There is, for example, only one way a slide ruler can be manipulated, so they are not general-purpose machines. By contrast, digital computational artifacts, particularly modern computers, have a much more complicated structure-function relation. Modern computers are not designed to perform only one type of computation, but are, in James H. Moor's words:

"logically malleable in that they can be shaped and moulded to do any activity that can be characterised in terms of inputs, outputs, and connecting logical operations...The potential applications of computer technology appear limitless. The computer is the nearest thing we have to a universal tool" (Moor 1985, p. 269).

In the previous chapter (section 2.3), I have argued that physical structures are multiply usable. This is particularly true for digital computers[33], because they are logically and functionally malleable and are therefore the nearest thing we have to a universal tool (see also Floridi 1999). In analog computers, there is usually a one-to-one relation between structure and cognitive function, i.e., a particular physical structure can only perform one kind of function, e.g., mathematical computations. By contrast, digital computers exhibit a one-to-many relation between structure and cognitive function, i.e., a particular physical structure can perform many kinds of functions. Their functional malleability comes from the fact that they can be re-programmed and moulded to the users' needs, but also because they are a medium in which a variety of representational systems can be expressed and manipulated. The information that computers exhibit on their screens are highly dynamic and malleable and can be iconic (e.g. pictures), indexical (e.g. real-time weather radar), or symbolic (e.g. language).

[33] I am talking here about modern digital computers such as desktops, laptops, tablets, and smartphones. The first digital computers were not general-purposes machines, but typically only performed one type of computation.

Computers also exhibit other information such as programming languages and software programs. These are typically symbolic, as they acquire their meaning from logical rules, but from a user-centered perspective, these do not really matter as they mainly happen inside the computer and are more relevant for computer programmers and software developers than for users. From a phenomenological user-centered perspective, what matters is what happens on the screen, not inside the computer, because cognitive functions supervene on information which are only visible on the screen.

Finally, these three kinds of relations between physical and representational structure (i.e. adding, constituting, and computing) are not mutually exclusive, but quite often overlap. We have seen that a token artifact may display a combination of iconic, indexical or symbolic properties. It may, likewise, also display a combination of additive, constitutive, and computational properties. For example, a scale model may have paint added onto its physical structure (additive and constitutive), a thermometer has a static representational structure and a dynamic constitutive structure (additive and constitutive), or a computer can simulate a scale model (constitutive and computational). Consequently, a useful way of conceptualizing this trichotomy of relations between physical and representational structure is by seeing a token representational artifact as predominantly additive, constitutive, and computational.

3.5 Misrepresentation and Malfunction

Having looked at structure-function relations, let us now go back to the content of the representations, i.e., their information. Representations can misrepresent their content: they can be incorrect, inconsistent, or inaccurate, which, depending on the degree of misrepresentation, will hinder cognitive performance. When an external representation misrepresents its content, it can be said to malfunction, because it is not doing what it is selected to do. Misrepresentation and thus malfunction can happen on different levels: when the physical structure breaks down, or when the representation itself is incorrect, inconsistent, or inaccurate. In the previous subsection, I pointed out that there are at least three relations between physical and representational structure. Representational structures can be (1) carried by or added onto a physical structure, e.g., a map printed on paper; (2) statically or dynamically constituted by physical structures, e.g., a scale model or thermometer; and (3) computed by a physical structure, e.g. slide ruler or digital computer. In these distinct (but sometimes overlapping) relations, misrepresentation and malfunction can go as follows.

In all these relations, representations depend for their existence on a physical substrate. When that substrate, for whatever reason, breaks down, the representations no longer have a substrate and will also breakdown or deteriorate. Because representations depend on their substrate in different ways, their breakdown can also happen in different ways. First, the physical structure (e.g. paper) onto which a representation (e.g. map of Manhattan) is added may simply wear out, get damaged, may be poorly designed, et cetera. When this happens one can no longer interpret the representations and it therefore becomes malfunctional. Second, the physical structure that statically or dynamically constitutes a representational structure can breakdown: the wood of a scale model may shrink or expand due to temperature and/or humidity differences and no longer accurately represent the ratios of the structure it represents, or mercury can leak out its column, resulting in a thermometer that displays a lower temperature than it in fact is. Third, analog or digital computational structures can breakdown: the slide of a slide ruler may break in two pieces, or the transistors and circuitry of a computer can short circuit or overheat, resulting in a malfunctioning, but not necessarily *mis*representing artifact. When digital computational artifacts breakdown, they usually lose their capacity to represent altogether, thereby becoming *non*-representing artifacts.

Things can also go wrong purely at the representational level, in which case the physical structure of the artifact is functioning properly, but the representations it exhibits are incorrect, inconsistent, or inaccurate. This can happen for various reasons: a map can misrepresent topographical structures, because its designer accidentally used the wrong colours to indicate land relief. A diagram can misrepresent the amount of carbon dioxide in the earth's atmosphere, because the measurements on which it is based were wrong. A clock can misrepresent the time, because its user mistook an eight for a zero when setting the time. A scale model or blueprint can misrepresent its content, because the architect was given incorrect information by its client. Or a slide ruler can misrepresent the answer to a calculation, because its logarithmic scale is inaccurate due to a manufacturing mistake. In these examples, misrepresentation happens purely due to their representational structure, not their physical structure, which may be caused by designers, manufacturers, users, or other agents. Note that, in some cases, misrepresentation can be seen as a continuum. A map can misrepresent topographical information, but accurately represent geographical structures and can still be used for

navigating but not for topographical purposes. So it may be malfunctional for one purpose, but functional for another.

In the above examples, representations (partly) misrepresent their content because they are incorrect, inconsistent, and/or inaccurate. They can also be outdated, which is not the same as misrepresentation or malfunction. Consider the following example: a map of New York City (NYC) from 1920[34] is currently outdated, as it no longer accurately represents the current structure of NYC. Intuitively, it may seem as if the map is misrepresenting and thus malfunctioning, because it does not accurately represent the current structure of NYC and can therefore not be used for navigation. However, it does what it is supposed to do, i.e., it is performing its proper function, which is to accurately represent NYC's geographical structure of 1920. The map was never designed nor selected to be used for navigational purposes almost a hundred years after it has been made. So, strictly speaking, one cannot say it is misrepresenting, because it represents what it is supposed to represent. Moreover, although it may not be useful for navigation, this is not to say that it is useless. Historians may find such a map helpful for their inquiries. Most maps are typically not designed for historical enquiries, but historians can, of course, select maps for their idiosyncratic research purposes, thereby attributing a cognitive system function to it.

A distinction between representational and cognitive function is helpful here. A representational function of a cognitive artifact is to accurately represent what it is selected to represent and a cognitive function is either a proper or system function of a cognitive artifact. Consider (again) a NYC-map from 1920. The representational function of the map is to accurately represent the geographical structure of NYC of 1920. In 1920, the purpose for which such a map was deployed was, in most cases, to navigate. But the same map can currently be used by an urban historian to study the structure of the city, not for navigational purposes, but for historical purposes (whatever they may be). The representational function of the map is in both cases the same, i.e. to represent the structure of New York City in 1920, but the cognitive purpose for which the map is deployed is different. Hence, representational functions and cognitive functions can be seen as distinct. So, in short, the representational

[34] That is, before urban planner Robert Moses dramatically and quite controversially changed the city's structure.

function of a cognitive artifact is the answer to the question: what information does the artifact represent? And the cognitive function is the answer to the question: what does an agent do with that information?

3.6 Interpretation and Consumption

In order to establish a genuinely triadic relationship, an agent has to interpret or consume the representation. For this to happen, an agent has to understand the dyadic relationship between the vehicle and its target. An agent thus has to understand the representation *qua* representation. This means that an interpreter has to realize that icons display isomorphic relations with their targets, that indices have causal relations with their targets, and that symbols are based on rules and conventions. So when I look at a map of the London subway system, I understand that I am looking at a predominantly iconic representation. Due to the iconic properties of the vehicle, I understand that the London subway system is an existing structure in the world and because of this understanding I am able to form beliefs about certain properties of the London subway system, e.g., the order of some of the stations. Thus representations are psychologically efficacious in that they cause beliefs about the informational content of the vehicle. In this sense, they causally and informationally mediate between agent and world.

Representational vehicles can be misinterpreted, which, depending on the degree of misinterpretation, will hinder cognitive performance. This occurs when the vehicle accurately captures its target, but for whatever reason, the agent misinterprets the vehicle and attributes certain properties to its target that do in fact not exist, resulting in false beliefs about aspects of its target. In such cases, there is no genuinely triadic relationship between agent, vehicle, and world, but a mere dyadic relationship between agent and vehicle. It is possible that an agent only misinterprets certain elements of the vehicle. For example, when navigating the London subway system with the aid of a map (not an easy cognitive task for a tourist), I may correctly interpret my current location on the map, say, Tower Hill, but misinterpret how many stations it is from Tower Hill to Piccadilly Circus and, consequently, form a false belief about its location. In this example, there is a mixture of a true and false beliefs, but due to the true belief, a genuinely triadic relation is established.

There can be many reasons for misinterpretation: it can be due to the agent (e.g. poor eye-sight or being tired), due to environmental aspects (e.g. poor light or noisy), or due to the design or properties of the representation (e.g. too much irrelevant information which distracts from interpreting the relevant parts). In order to prevent misinterpretation, it is important that representations are designed such that they only contain information relevant for a particular task (see Norman 1991, 1992, 1993, 1998; Rogers 2004). Simplicity is power when it comes to the design of external representations.

4. Ecological Cognitive Artifacts

Having outlined the genus of representational cognitive artifacts, I now focus on the genus of ecological cognitive artifacts. These are characterized by exhibiting non-representational or ecological information. To give a brief example: when playing Tetris, the artificial rotation of a zoid clearly has a functional role in performing a cognitive task (Kirsh & Maglio 1994). This functional role, however, is not established by exhibiting representational properties. Zoids do not stand in for something else, they have no representational content, and do not mediate between an agent and a target. They are abstract geometrical shapes that are interacted with directly, without them representing something outside the game. So, rather than triadic situated cognitive systems (agent-representational artifact-target), ecological artifacts concern dyadic situated cognitive systems (agent-artifact). The main difference between these two kinds of systems is that there is no representational target in case of dyadic systems. The functional role of the artifactual elements in such dyadic systems is thus established through non-representational or ecological information. But what exactly is ecological information? How can it aid their users in performing a cognitive task? And how can it be categorized?

The work of cognitive scientist David Kirsh (1995, 1998, 2006, 2009a; Kirsh & Maglio 1994; Hollan, Hutchins & Kirsh 2000) is a useful starting point in answering these questions. In his paper, *The Intelligent Use of Space*, Kirsh (1995) makes a tripartite distinction between spatial arrangements that simplify choice, perception, or internal computation. Thus the cognitive function of such spatial arrangements is not representational, but to simplify decision-making, perception and recognition, and internal cognitive processing. Although this a helpful and insightful categorization, it is

(like all other categorizations discussed in this thesis) cognition-centered, i.e., it starts with a cognitive agent and then categorizes artifacts on the basis of the cognitive processes to which they contribute. In the taxonomy I am developing in this chapter, I take an information-centered approach and taxonomize cognitive artifacts on the basis of their informational properties. With this information-centered approach in mind, I recycle and reclassify some of Kirsh's examples and focus on their informational properties, rather than on the cognitive processes to which they contribute. By doing so, I distinguish between two species of ecological artifacts, those that obtain their function in virtue of physical-spatial structures and those that obtain their function in virtue of manipulable physical structures.

4.1 Spatial Ecological Cognitive Artifacts

Human agents quite frequently make use of space for cognitive purposes (see also Hutchins 2005; Knappett 2010; Woelert 2010). This is so commonplace, Kirsh argues, that "we should not assume that such cognitive or informational structuring is not taking place all the time" (Kirsh 1995, p. 33). The intelligent use of space enables us to encode important information into artifacts that are typically neither designed nor selected for cognitive purposes and thus mainly (though not solely) concerns improvised or system uses of artifacts. Some straightforward examples include consistently leaving car keys on a certain spot in your apartment so that you know where they are, putting an article you have to read on top of the pile on your desk, leaving a book open and turned upside down so that you know where you have stopped reading, tying a string around your finger as a reminder, or leaving a rented DVD on your desk as a prompt to bring it back to the video store. By putting artifacts in specific locations that are either deliberately usual or deliberately unusual, we encode information into the artifact and its location, such that when we encounter the artifact in that location it typically prompts a memory. Such artifacts may be referred to as spatial ecological cognitive artifacts.

Consider another example: when doing the dishes, it is not always clear which items have been washed and which ones have not. In order to keep track of the items that have been washed, it is helpful to create spatial categories of "washed items" and "unwashed items" by putting them in certain locations. Most kitchen sinks have designated areas for items that have been recently washed. These areas are not just practical, so that residual dishwater can drip away without spilling it on the kitchen

counter, but also have cognitive functions, as they simplify perception and reduce memory load in a task. Whether this concerns a cognitive proper function or cognitive system function is a question that needs empirical evidence for a definitive answer. Designers most likely did not design that area so that kitchen users can deploy it to ease their perception and memory. Neither did kitchen users deliberately select a kitchen with such an area to ease their perception and memory. Such an area, I think, is most likely designed and selected to put items that have just been washed so that residual dishwater can drip away, so its proper function is pragmatic rather than cognitive. When designing or selecting kitchens, designers and users are probably not aware of the cognitive function such an area can have. Only during improvisation it becomes clear that we may use the space in our kitchen for cognitive purposes and therefore it is best seen as a system function.

Kirsh (1995) describes a more idiosyncratic case of someone who is dismantling a bicycle and then puts certain parts on a sheet of newspaper placed on the floor. The newspaper demarcates a spatial boundary within which certain items are placed, in that way structuring the task space and making items easier to locate[35]. The user may also place the dismantled items in such a way that, when reassembling the bicycle, the items that need to be reassembled first are located closer to the user than the items that are reassembled later in the process. "The virtue of spatially decomposing the task is that one need not consult a plan, except at the very highest level, to know what to do. Each task context affords only certain possibilities for action" (Kirsh 1995, p. 44). Kirsh here is referring to a mental plan, but it could also be an external one, e.g., an assembly guide, manual, or blueprint of some kind. There is no need for a manual (or an elaborate mental plan) if all the parts are placed such that they correspond to the correct order of actions for reassembling the bicycle. Spatial structuring of artifacts thus makes both complex internal representations and external representations superfluous. It is much more efficient to spatially structure the artifactual task environment such that it affords the most efficient and environmentally-embedded plan.

[35] This is an interesting example, I think, because a proper cognitive artifact (i.e. a newspaper) now serves a cognitive system function. Both its proper and system function are cognitive in nature, i.e., to inform its reader about events in the world, or to demarcate a spatial boundary, in that way making a perceptual task easier to perform.

The intelligent use of space, however, is not restricted to physical or actual space, as it also frequently occurs in virtual space. In connection to Human-Computer Interaction, Hollan, Hutchins, and Kirsh provide some suitable examples. Computer users, they write;

"Leave certain portals open to remind them of potentially useful information or to keep changes nicely visualized; they shift objects in size to emphasize their relative importance; and they move collections of things in and out of their primary workspace when they want to keep certain information around but have other concerns that are more pressing" (Hollan, Hutchins & Kirsh 2000, p. 190).

So the way we spatially organize the items and structures on our screen helps us in performing certain computer-related tasks. We may, for example, prioritize certain information by leaving portals open or by making them larger than other portals. We may also organize the items on our desktop or in our navigation menu's such that they reflect their importance. Often used items typically inhabit a more prominent position than items that are used less often (for example, by placing them in preferred positions on one's desktop or menu) thereby making them easier to locate (see also Heersmink 2013a). This is most likely why computer interfaces afford such options, so they can be seen as proper functions of computer systems. These everyday examples show that we organize artifactual elements within space for cognitive purposes, thereby encoding important information into physical-spatial structures.

4.2 Dynamic Ecological Cognitive Artifacts

Some ecological artifacts obtain their cognitive function in virtue of their *manipulable* physical structure. During or after manipulation of such artifacts, new information emerges that is important to aid performing a cognitive task. For example, when rearranging letter tiles in Scrabble to prompt word recall, new information emerges from their spatial configurations (Kirsh 2009a). In this case, the non-representational information of the tiles (i.e. individual letters), their spatial location in relation to the other tiles, and the words and openings on the board are important for performing the task. Novel and larger informational units emerge when the tiles are rearranged, which (ideally) prompts the recall of words with as many letters as possible and that fit into existing letter structures on the board.

Kirsh points out that one of the purposes for rearranging letters tile, is to generate and draw attention to often occurring two- and three-letter combinations that figure in words such as, for instance, "ES", "TH", "IN", "REA", et cetera. Note that such information (i.e. letter combinations) is often not representational in nature, although there is a symbolic element to it, because there are logical rules and conventions for creating words out of letters. So perhaps letter tiles can be seen as sub-representational, as they are the building blocks of proper representations. The point here is that the cognitive function of the artifacts supervenes on their manipulable physical structures and by manipulating their physical structure one automatically manipulates their information-structure, too. Whether this concerns a proper or system function is again a difficult question to answer. I am not sure whether designers or users were aware of this particular function when they, respectively, designed or selected the game. This particular example of the intelligent use of space is, I think, most likely a system function.

Virtual structures, too, can be manipulated such that new information emerges from their spatial configurations. When playing Tetris, one has to rotate a geometrical shape called a "zoid" so that it fits into a specific socket in the lower regions of a geometrical template. One can either choose to mentally rotate the zoid or to rotate it by means of a button-push. Experienced players have learned to rotate the zoid by means of a button-push, not only because it is significantly faster, but also because it relieves the brain from performing mental rotation (Kirsh & Maglio 1994; see also Clark & Chalmers 1998). Their functional role is established due to a constantly manipulable virtual structure, i.e., *because* the zoids can be manipulated, a user delegates rotation to the computing device, in that way enabling a user to decide quicker whether or not it fits into a socket in the lower regions of the task space. Like with rearranging Scrabble tiles, new task-relevant information emerges only in relation to some other structure. In Scrabble, new information emerges when two or more tiles are positioned in a certain way. In Tetris, new information emerges when the zoid is spatially orientated in relation to a template in a certain way. The spatial orientation of the zoid in itself, i.e., without taking into account the structure of the template in which it has to fit, is not sufficient for performing the task. Both zoid and template are important for generating the task-relevant information. Finally, the above artifacts are much more dynamic than cases of spatial ecological cognitive artifacts. They are not about developing static

spatial categories in which artifacts are placed, but about dynamic and constantly changing information in an ongoing task[36].

4.3 Structure-Function Relations

The cognitive function of ecological artifacts supervenes on ecological information. They therefore have different structure-function relations as compared to representational artifacts because they lack a representational layer. How to conceptualize the relationship between the physical structure of the artifact, its spatial properties or location, and its ecological information? In other words, how does ecological information emerge from physical-spatial configurations? There are at least two different kinds of relationships between physical structure and location. The first concerns spatial ecological artifacts and the second dynamic ecological artifacts.

First, when physical structures are intentionally placed at a particular location, a hybrid of a physical-spatial structure comes into being, from which task-relevant information emerges (see figure 3). Car keys, for example, need to be intentionally placed at a particular location in order to function as a memory aid. If, for some reason, an agent forgets to place the car keys at their usual location, then the task-relevant information is not encoded into a physical-spatial structure and, consequently, there is no cognitive function. This will significantly hinder cognitive performance, as an agent now has to look for one's keys. In this case, the information encoded into a physical-spatial structure is self-referential, i.e., the location of the car keys is *identical to* its information-structure. Likewise, leaving a book open and turned upside down so that you know where you have stopped reading, contains a self-referential information-structure, i.e., the information that that particular physical-spatial configuration contains, refers to itself. To better understand self-reference, it is helpful to briefly compare it to other spatial ecological cognitive artifacts that do not exhibit self-reference. When I leave a rented DVD on my desk as a prompt to bring it back to the video store, put an article on top of the pile as a reminder to read it, or tie a string around my finger as a reminder for some action, the artifacts are not self-referential because their information refer to an action, not to themselves.

[36] Elsewhere (Heersmink 2013b) I referred to this kind of artifacts as "structural ecological cognitive artifacts". Given that their function supervenes on constantly manipulable physical structures I think the word dynamic is more appropriate now.

In some cases, the physical structure of the artifact is necessary for establishing its proper function (e.g. to start a car), but not for establishing its cognitive system function, for which its location is much more important. In case of car keys, for example, the key can be broken, in which case it can no longer perform its proper function, but if it is placed at its usual location, it can still perform its cognitive system function. In other words, even malfunctioning car keys placed at the correct location, exhibit the relevant information-structure, demonstrating that location is fundamental for establishing such cognitive system functions. Physical structure seems to be secondary in such cases.

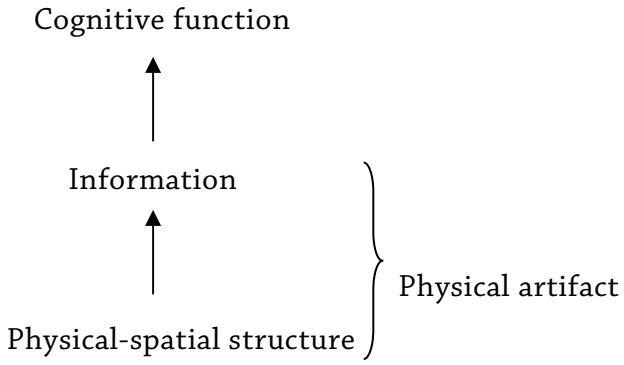

Figure 3. Relation between physical-spatial structure, informational structure, and cognitive function.

Second, information-carrying structures can be manipulated such that new information emerges (see figure 4). This happens with letters tiles in Scrabble and zoids in Tetris. In case of letter tiles, the information-structure (i.e. letters) is carried by a physical structure in a very similar way as a piece of paper carries a map. In case of zoids, their virtual structure is their information-structure, i.e., their form and shape constitute their information-structure in a very similar way as a scale model constitutes its information-structure. In either case, by manipulating those structures one automatically manipulates the information they contain, too, which establishes their cognitive function. Also, their particular information-structure emerges only when two or more letters tiles are organized in a particular way, or, in case of zoids, relevant information emerges only in relation to a template in the lower regions of the task space. So it may be argued that relational constructs or hybrids between two or more artifacts (e.g., two letter tiles or the zoid and the template) emerge only when they are orientated in a particular way.

In case of spatial ecological artifacts, a physical-spatial hybrid was created by intentionally placing an artifact or artifacts at a particular location. From this physical-spatial hybrid, task-relevant information emerges when an agent, at some later point, sees the artifact in that location. In case of dynamic ecological artifacts, physical structures are manipulated in relation to some other structure. Rotating zoids is helpful only because it allows a user to assess whether its new orientation fits some template in the lower regions of the task space. What matters here is thus the spatial relation between zoid and template. It is thus not about an artifact that has been put at some location, but about dynamically changing the spatial relation between two (or more) artifacts (e.g., a zoid and template).

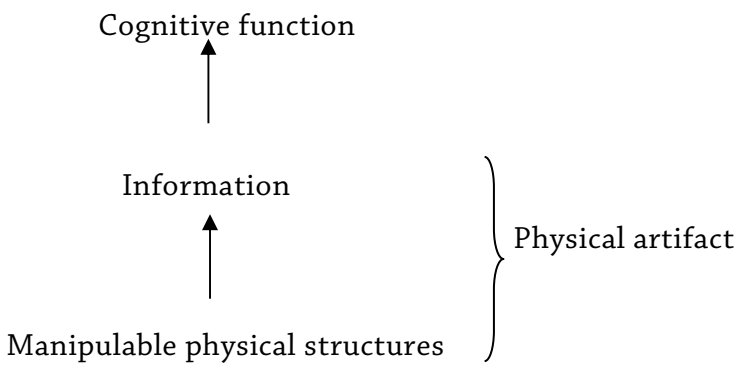

Figure 4. Relation between manipulable structures, informational structure, and cognitive function.

4.4 Representational and Ecological Artifacts

Ecological artifacts, I think, are instances of what roboticist Rodney Brooks (1999, 2002) calls "using the world as its own best model" and what Clark (1989) calls the "007 principle". The point of these notions is: why create an expensive internal representation of the world, when you can use the world itself as a model? In a similar way we may ask: why create an expensive external representation of the world, when you can use the world itself to interact with and use it to facilitate your cognitive tasks? One could, for example, consult a manual for how to reassemble a dismantled bicycle. But one can also structure the items such that they facilitate the reassembling process, in that way streamlining the task and making a cost-expensive external representation superfluous. In Clark's words, "evolved creatures will neither store nor process information in costly ways when they can use the structure of the environment and

their operations upon it as a convenient stand-in for the information-processing operations concerned" (Clark 1989, p. 64).

For analytical purposes, I presented representational and ecological artifacts as distinct genera. Up to this point, a reader may get the impression that a cognitive artifact either exhibits representational or ecological information. This analytical distinction was helpful in that it allowed me to emphasize and conceptualize their distinct cognition-aiding informational properties. However, in some cases, cognitive artifacts exhibit a combination of representational and ecological properties. To give an example: in a bookcase in which the books are alphabetically organized, representational and ecological structures jointly facilitate a perceptual task. The representational structures are the names and titles on the back of the books and the ecological properties are the spatial order in which they have been placed. In this example, the representational properties determine the spatial structure, which in turn, supports the representational structure. So alphabetically organized books may be seen as a predominantly representational cognitive artifact. Thus the genera of representational and ecological artifacts are not mutually exclusive, but in most cases either one of those properties is predominant.

In the second chapter, a number of examples of cognitive artifacts that have both representational and ecological properties have been given. Recall that Donald, Norman, and Clark argued that some artifacts exhibit malleable representational structures that afford adding, erasing, copying, restructuring, reformatting, or otherwise manipulating external representations. It is partly, or perhaps even largely, due to their *representational malleability*, that these artifacts afford and facilitate higher-order cognition. This happens, for example, when writing an academic paper, making a difficult calculation with pen, paper, and numerals, and preliminary sketching. In these cases, the artifactual elements do not have a functional role in performing a cognitive task solely because they have representational properties, but also because they have spatial and malleable properties. This representational malleability is very similar to the malleability of letter tiles in playing Scrabble, or the malleability of zoids in playing Tetris. Although their information is non-representational, the malleability principle is largely the same.

5. Concluding Summary

This chapter developed a taxonomy of the artifactual element in situated cognitive systems by drawing on artifact categorization in archaeology. In developing this taxonomy, an information-centered approach was taken, i.e., I took as my point of departure the specific informational properties of cognitive artifacts and then taxonomized them on the basis of those properties. Current categorizations focus on representational artifacts and thus neglect non-representational or ecological artifacts. They also tend to focus on proper cognitive artifacts and often overlook system cognitive artifacts (Kirsh being a notable exception). These categorizations therefore have a significantly smaller target domain. Moreover, all current categorizations are cognition-centered, i.e., they start with human cognition and then work their way towards the (proper) representational artifacts that agents deploy to aid their cognition and realize their cognitive purposes. My information-centered approach is a valuable alternative to cognition-centered approaches, as it results in a much richer and detailed taxonomy.

In the developed taxonomy, three levels or taxa are distinguished, those of family, genus, and species. The family includes all cognitive artifacts (i.e. proper, system, representational, and ecological) without further specifying functional or informational properties. On the second level in the taxonomy, I distinguish between two genera: representational and ecological cognitive artifacts. On the third level, these two genera are further divided into species. In case of representational artifacts, those species are: iconic, indexical, and symbolic. In case of ecological artifacts, those species are: spatial and dynamic. Within species, I identified type cognitive artifacts. The categories in the taxonomy are not mutually exclusive, but one particular informational property is usually predominant. See figure 1 above for a diagram of this taxonomy.

In addition to taxonomizing cognitive artifacts, I also conceptualized one of their fundamental metaphysical properties, i.e., their structure-function relations. Cognitive artifacts are particularly interesting in this regard because they do not have a dual, but triple nature, which means that their function cannot straightforwardly be related to their physical structure, as there is an additional representational or informational layer in between their structure and function. Representational structures can be (1) carried by or added onto a physical structure, e.g., a map printed on paper; (2) statically

or dynamically constituted by physical structures, e.g., a scale model or thermometer; and (3) computed by a physical structure, e.g. slide ruler or digital computer. Ecological artifacts lack representational properties and therefore have different structure-function relations. In such artifacts, information can be encoded as follows: (1) physical structures can be intentionally placed at a particular location, thereby creating a hybrid of a physical-spatial structure, from which task-relevant information emerges; or (2) information-carrying structures can be manipulated such that new information emerges from their new configurations.

Part II

Situated Cognitive Systems

5

The Complementary Integration of Agents and Artifacts

1. Introduction

Part I of this thesis covered the categories and metaphysics of the artifactual element in situated cognitive systems, without being too concerned with the agents using them. In Part II, the focus will shift and I will look at situated *systems* by conceptualizing how cognitive artifacts and their users are integrated into larger systems. One useful way to look at situated cognitive systems is through the lens of distributed and extended cognition theory (Hutchins 1995; Kirsh 2006; Menary 2007a; Clark 2008b; Sutton 2006, 2010). Extended cognition theory claims that human thought is, under certain conditions, distributed across an embodied agent and cognitive artifacts or other external resources. Such resources are then not seen as merely external aids or scaffolds for thinking, but are sometimes proper parts of a distributed or extended cognitive process. Cognitive states and processes are thus conceptualized as hybrids or amalgamations of neurological, bodily, and environmental objects and processes (Wilson 1994, 2004; Rowlands 1999, 2009, 2010). On this view, some of the vehicles of cognitive processes are sometimes not located in the brain, but in the cultural and artifactual environment. It is thus a metaphysical claim about the location of human cognition, implying significant methodological and epistemological consequences for the philosophical and scientific study of human thought.

1.1 Parity and Complementarity

Sutton (2010) has identified two distinct, but historically overlapping, waves in extended cognition theory. The first wave is mostly based on the parity principle and is advocated by Clark and Chalmers (1998), Mike Wheeler (2010, 2011), and others. The parity principle is as follows:

"If, 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 & Chalmers 1998, p. 8).

The parity principle focuses on functional isomorphism between internal states and processes and external states and processes. By stressing functional isomorphism, it downplays differences between internal and external states and processes, in that way implying that the specific properties of cognitive artifacts are not important for better understanding situated cognitive systems. As long as there is a relevant (and rather generic) similarity with internal states or processes, then we do not need to pay attention to the specific properties of the external resource. However, the particular properties of the external resource are often important for better understanding how the overall system performs a cognitive task (Norman 1993; Zhang 1997; Heersmink 2013b).

Stressing isomorphism is, moreover, not fruitful as there are differences between the internal and external. In the case of memory, for example, there are differences in how internal memories and external memories are stored and processed. Internal memories are stored in neural networks that are subject to blending and interfering, which means they are shaped and updated according to other previously stored and new incoming information. In contrast, external memories are stored in discrete iconic or symbolic format that are static, are less dynamic, and are not automatically integrated with other information (Sutton 2010). There are thus differences in the informational properties of internal and external memories.

These and other drawbacks of the parity principle have led some theorists to argue for a different route to extended minds. This route is based on the complementarity

principle, which downplays functional isomorphism and emphasizes complementarity between the internal and the external. On such a view, cognitive artifacts need not be similar to internal states and processes, but often complement internal states and processes with *different* properties and functions. In fact, complementing brain functions is often the point of deploying cognitive artifacts: so that they can perform functions the brain cannot do or cannot do well. There is often no point in externally replicating what the brain can already do. Jointly, brain-plus-artifact is a much more versatile and powerful problem-solving system than an embodied brain alone. On a complementarity view, then, we should study the variety of situated cognitive systems by focussing on the particular properties of both embodied agents and cognitive artifacts (Sutton 2010) and by focussing on the functional integration of the two (Menary 2007, 2010b, 2010c). Rob Wilson and Clark defend a similar view.

"Contrary to any requirement of fine-grained similarity then, what the friends of extended cognition actually expect, and study, are hybrid processes in which the inner and the outer contributions are typically highly distinct in nature, yet deeply integrated and complementary" (Wilson & Clark, 2009, p. 72).

Wilson and Clark thus claim that internal and external parts of situated cognitive systems are deeply integrated and exhibit complementary properties. They further point out that "Tracing and understanding such deep complementarity is, we claim, the single most important task confronting the study of situated cognition" (Wilson & Clark, 2009, p. 70). However, what they do not point out is how and how deeply the internal and external components are integrated, i.e., they do not look at the process of integration. Given that the complementary properties and functions of cognitive artifacts and other external resources are integrated into the onboard cognitive system to varying degrees, it is important to have the tools to conceptualize the degree and varieties of integration. The goal of this chapter is therefore to develop some of those tools by proposing and further developing a number of dimensions along which to conceptualize integration.

1.2 Method, Motivation and Organisation

Sutton (2006; see also Sutton et al 2010), Wilson & Clark (2009), Kim Sterelny (2010), and Richard Menary (2010c) have articulated the idea of a dimensional analysis of the coupling between agents and external resources. By synthesising and building on their

work, I develop a multidimensional framework for conceptualizing the different kinds of cognitive interactions and complementary integration between agents and artifacts. This framework consists of the following dimensions: epistemic action and information flow, speed of information flow, reliability, durability, trust, procedural transparency, informational transparency, individualization, and transformation (see Heersmink 2012b for a preliminary version of this framework). These dimensions are all matters of degree and jointly they constitute a multidimensional space in which situated cognitive systems can be located. This is true not only of those cognitive systems that are or may be extended or distributed, but also of those that are not. The higher a system scores on these dimensions, the more functional integration occurs, and the more tightly coupled the system is.

Importantly, although motivated by complementarity-based extended cognition theory, this framework is not restricted to the extended mind cases in which a minimal requirement is a two-way interaction. Clark and Chalmers (1998) have characterized extended minds as "coupled systems" in which there is a two-way interaction between an agent and artifact. In such coupled systems, agent and artifact both play an active causal role in an overall cognitive process. Two-way interaction is thus an important element for cognitive extension. But, it is still important to better understand one-way or monocausal relations such as, for example, navigating with the aid of traffic signs, looking up the departure time of a train in a timetable, or assembling a piece of furniture with the aid of a manual. Indeed, a high proportion of cognitive artifacts have a monocausal influence on human thought and behaviour, so for explanatory reasons (i.e., explanatory scope and completeness), it would be unwise to exclude monocausal relations from the picture even if these cases are not candidates for extended cognition. In order to develop a more inclusive picture, I develop a multidimensional framework in which most of the dimensions are also helpful for better understanding monocausal interactions.

There are at least three reasons why we need such a multidimensional framework. First, because it encourages and provides a toolbox for a detailed study of conceptual and empirical cases of situated cognitive systems, in that way providing us with a much needed and richer understanding of the variety of complementary relationships between agents and artifacts. This toolbox is not only helpful for philosophers interested in situated cognition, but also for psychologists, anthropologists, Human-

Computer Interaction theorists, design theorists, media theorists, and others interested in human-technology relations. It may guide empirical research by providing the dimensions along which to study and observe agent-artifact interactions[37] and may potentially result in improved interface designs. Although this toolbox is developed primarily to analyse agent-artifact relations, it should be equally useful to analyse agent-agent systems. Conceptual and empirical research on socially embedded or distributed cognitive systems (e.g. Wilson 2005; Barnier, Sutton, Harris & Wilson 2008; Theiner 2013; Tollefson, Dale & Paxton 2013) would thus also benefit from this framework.

Second, conceiving of situated cognitive systems, including those that are extended and distributed, in terms of dimensions that are matters of degree, provides a much more realistic view of such systems. Seeing situated systems as either embedded or extended is not a particularly fruitful way of conceptualizing such systems, as some may be more embedded or extended than others (Sutton et al 2010). The nature of the functional and informational distribution and integration differs in each particular case. So rather than providing a set of necessary and sufficient conditions, it provides a toolbox for investigating the degree and nature of distribution and integration.

Third, using artifacts for *cognitive* purposes is arguably one of the most sophisticated forms of tool-use and seems to be distinctively human (see also Vaesen 2012). If cognitive artifacts make us smart and uniquely human by complementing our limited cognitive abilities, it is important to have a framework that gives us a richer and deeper understanding of the interactions with such artifacts. It is much smarter to complement our shortcomings, rather than to replicate existing onboard capacities. We seem to have evolved to incorporate tools into our cognitive (and bodily and perceptual) systems and to better understand this capacity it is important to better understand the variety of relationships that are established between us and our cognitive artifacts.

The chapter is organised as follows. It starts by discussing the complementarity principle (section 2). Next, the dimensions for conceptualizing the degree of complementary integration between agent and artifact are presented (section 3). Having developed this multidimensional framework, I investigate the relations between

[37] In chapter 7, I briefly outline how these dimensions may be used to guide empirical research.

some of the dimensions and look at how particular situated cognitive systems can shift to different regions on the multidimensional space (section 4). Finally, a concluding summary is given (section 5).

2. The Complementarity Principle

Given the conceptual and methodological consequences of the parity principle mentioned in the introduction (see also Rupert 2004; Walter 2010), Sutton has identified and articulated a distinctive route to extended cognition based on what he refers to as the complementarity principle:

"In extended cognitive systems, external states and processes need not mimic or replicate the formats, dynamics or functions of inner states and processes. Rather, different components of the overall (enduring or temporary) system can play quite different roles and have different properties while coupling in collective and complementary contributions to flexible thinking and acting" (Sutton 2010, p. 194).

This principle downplays functional isomorphism between inner and outer states and processes and argues for complementary properties and functions of cognitive artifacts or other external resources. Human brains are not good at storing large chunks of information for a long period of time in discrete manner and therefore offload information onto the environment that supplements our internal memory and information-processing capacities. "Biological traces are typically integrative, active, and reconstructive, but in using them we hook up to more endurable and transmissible exograms, mostly of our own making, which supplement and extend our powers" (Sutton 2010, p. 205). So external information systems exhibit different properties as compared to internal information systems, in that way complimenting biological memory. Cognitive artifacts, however, do not just complement our memory, but a variety of cognitive capacities. Consider three brief examples.

First, we have a cognitive capacity to roughly estimate the quantity of certain aspects of our environment such as, for instance, wind direction, or the weight and size of objects. However, when using devices such as weather vanes, scales, and rulers, we can measure and quantify those aspects with a much higher accuracy and, in that sense, they complement existing perceptual capacities to estimate quantities. Moreover, other

devices such as compasses, voltmeters, spectrometers, and barometers allow us to measure and quantify aspects of the world that we are unable to estimate or measure without those devices, because our sensory systems do not have access to those aspects of the world, i.e., we cannot perceive and quantify magnetic fields, volts, or ultraviolet light without artifacts. By exhibiting those functionalities, those artifacts add completely new capabilities to our perceptual and cognitive repertoire, thereby complementing our ability to map and quantity (properties of) the world.

Second, mental imagery is rather constrained in that it is relatively fixed, i.e., it is difficult to see new forms and components of a mental image, and it is limited, which means there is only so much we can imagine internally. By sketching, we offload information onto the environment, which may not only prompt and generate new forms and components, but also dramatically increases the amount of information an artist or designer can work with, as it provides much more detail and precision (Clark 2001; Tversky & Suma 2009). These properties of externalized information rather straightforwardly complement the constraints and limitations of our mental imagery, making us more powerful and creative designers and artists. Likewise, scientists often build and manipulate models of the target systems they are studying. They do so because is it difficult for them to store, let alone manipulate, complex scientific models in their head. Scientists need to create and manipulate external models whose functions and affordances complement the shortcomings of the pattern matching abilities of embodied brains. This has been referred as model-based reasoning (Magnani, Nersessian & Thagard 1999; Giere 2002a).

Third, internal reasoning processes have limited capacities, in part because we can perform only certain operations on internal information such as e.g. propositions, and in part because that internal information has a limited capacity in terms of the amount of information it can contain at any given moment. Thus when we develop an argument or line of thought, we can do only so much internally. But when using written language, e.g. with the aid of a word-processor, we can formulate much better and elaborate arguments or ideas. This is so because the representational malleability that word-processors afford, complements internal information-processing systems by allowing text to be erased, copied, restructured, reformatted, or otherwise manipulated. It thus allows us to perform operations on the offloaded information that are very hard, if not impossible, to perform in the brain. It also allows us to manipulate and work with

significantly larger amounts of information. These functionalities complement the reasoning capacities of embodied agents. So, despite Socrates' well-known worries about the effects of writing on memory, it allows us to significantly augment our onboard cognitive capacities.

These brief examples show that embodied agents deploy the functional and informational properties of cognitive artifacts to complement their onboard cognitive capacities. The human brain develops in an information-rich environment and has to learn how to complement its shortcomings by using the functionalities of artifacts to perform cognitive tasks in such a way that it does not waste expensive internal resources. The idea is that why perform a cost-expensive memory or information-processing task, when it can be delegated to an artifact?[38] A complementary integrated agent-artifact system is much more adaptable and powerful in problem-solving than an embodied agent without such artifacts. "Brains like ours need media, objects, and other people to function fully as minds" (Sutton 2010, p. 205). On a complementarity view, artifacts or other resources do not just augment existing brain functions by externally replicating them, but add new functionalities to existing ones by integrating them to varying degrees into a plastic neural architecture.

The complementarity principle does not have the structure of a conditional statement. Unlike the parity principle, it does not connect two conditional premises with an "if P, then Q" structure. Thus, it does not argue that if a cognitive artifact has a complementary role or function to inner processes or states, then it counts as a constitutive part of an extended cognitive process or system. Many things external to the brain somehow have complementary roles to internal cognitive states and processes. Formulating complementarity in terms of a conditional statement would leave it vulnerable to the cognitive bloat objection (Adams & Aizawa 2008). So rather than trying to stipulate what exactly demarcates an embedded system from an extended or distributed system, Sutton (2006, 2008, 2010) provides a taxonomy of the dimensions that are relevant for better understanding such systems. In the next section, we will look at these and other relevant dimensions.

[38] See also Clark's (1997) "007 principle" and Mark Rowlands' (1999) "Barking Dog principle".

3. Dimensions for Integration

In their article, *The Extended Mind*, Clark and Chalmers (1998) put forward a number of dimensions that describe the coupling between agents and cognitive artifacts. They argued that the dimensions of trust, reliability, accessibility, and past endorsement (as well as the parity principle) are important conditions for cognitive extension. However, as they point out, "The status of the fourth feature as a criterion for belief is arguable (perhaps one can acquire beliefs through subliminal perception, or through memory tampering?), but the first three features certainly play a crucial role" (Clark & Chalmers 1998, p. 20). Past endorsement as a condition for cognitive extension is thus questionable, but trust, reliability, and accessibility are crucial[39]. These three conditions are now referred to as "trust and glue". If one of these three conditions is not sufficiently satisfied, then cognition is not extended. So it provided a fairly clear set of criteria to distinguish between an external resource that is a proper part of an extended cognitive process and one that is a mere causal influence on an internal cognitive process.

A number of theorists (Sutton 2006; Sutton et al 2010; Wilson & Clark 2009; Sterelny 2010; Menary 2010c) further articulated the idea of a dimensional analysis of the cognitive relations between agents and external resources. Wilson and Clark identify two dimensions. First, the nature of the external resources, which may be natural, technological, or socio-cultural. Second, the durability and reliability of the overall situated cognitive system. Sutton et al (2010) take the dimensions of durability and reliability as well as the dimensions of "trust and glue" mentioned earlier by Clark and Chalmers (1998), and also briefly develop the dimension of transparency. Sterelny discusses three dimensions: trust, individualization, and individual versus collective use. And, finally, Menary focuses on the dimensions of manipulation and transformation.

All these dimensional approaches are perceptive and clearly help us in better understanding the cognitive relationship between agents and external resources, but they tend to overlook some dimensions. Moreover, it is not at all clear whether these

[39] For a critical discussion of these criteria, see (Rockwell 2010), and for a critical discussion of these criteria regarding extended memory, see (Michaelian 2012). For other analyses of the conditions for cognitive extension, see Rupert (2004), Roberts (2011), Adams & Maher (2012), Palermos (2012), and Smart (2012).

are all dimensions in the same sense of the word. Wilson and Clark (2009), for example, refer to the nature of the external resource (i.e. either natural, technological, or socio-cultural) as a dimension. Sutton et al (2010) and Sterelny (2010) refer to trust as a dimension. And Menary (2010c) refers to manipulation of external resources as a dimension. These theorists thus seem to use the notion of a dimension to indicate quite different aspects of situated cognitive systems.

In this section, therefore, I aim first to clarify the notion of a dimension. I then refine and synthesize some of the above mentioned dimensions into a coherent and systematic multidimensional framework, add a number of dimensions to the framework, and finally examine where and how some of these dimensions overlap and interact. Note that the framework I am proposing is not meant to be exhaustive, as there may be other dimensions relevant for better understanding situated cognitive systems. But compared to the above mentioned frameworks, the one developed in this chapter is more elaborate in that in contains more dimensions and is more integrative in that it examines how the dimensions relate to each other and interact.

It is helpful to clarify and distinguish a number of elements that are relevant for better understanding the underlying conceptual structure of each dimension. These elements are: (1) the cognitive profile or cognitive capacities of the human agent; (2) the informational and functional properties of the cognitive artifact; (3) the task environment and context of use; and (4) the kind of epistemic action and its cognitive purpose. Although they are essential parts of situated cognitive systems, these elements are not dimensions, but each dimension emerges out of the interplay between two or more of these elements. In any case, elements (1) and (2) are always essential, whereas elements (3) and (4) are, depending on each case, somewhat less essential. These dimensions are thus relational in the sense that they never depend on only one of those elements. So, in contrast to Wilson and Clark, who refer to the nature of external resources as a dimension, I use a different notion of dimension which is more in line with Sutton et al (2010) and Sterelny (2010), i.e., a dimension describes certain properties of the relation between an embodied agent and an external resource.

The first two dimensions (i.e., information flow and speed of information flow) are concerned with information flow and are therefore presented together. The other seven dimensions (i.e., reliability, durability, trust, procedural transparency, informational

transparency, individualization, and transformation) are presented in an arbitrary order. In section 4, the relation between the dimensions is further examined. Let us now turn to these dimensions.

3.1 Epistemic Action and Information Flow

The first of these dimensions are epistemic action and direction of information flow. It might seem as if these are two dimensions and analytically speaking they are, but they are so closely related that I treat them as a single dimension, as there is no information flow without epistemic action. Otherwise put, epistemic action is a necessary condition for information flow. Kirsh and Maglio (1994) have distinguished between pragmatic and epistemic actions (contrast Loader 2012). Pragmatic actions bring one physically closer to a goal (e.g. changing gears), whereas epistemic actions change the input to an agent's information-processing system and make mental computation easier, faster, or more reliable (e.g. deploying a map to navigate). Kirsh and Maglio write that

"We use the term epistemic action to designate a physical action whose primary function is to improve cognition by: (1) reducing the memory involved in mental computation, that is, space complexity; (2) reducing the number of steps involved in mental computation, that is, time complexity; (3) reducing the probability of error of mental computation, that is, unreliability" (Kirsh and Maglio 1994, p. 514).

To use the paradigm example: when playing Tetris, the strategy of advanced players is to push a button several times in order to quickly assess which position of the zoid fits a socket. Doing so, is much more efficient in terms of information-processing *speed* and information-processing *load* and therefore significantly simplifies the problem-solving task. Thus, pushing a button to change the position of a zoid is an epistemic action, because it changes the input to the computational system of the player and it makes mental computation (i.e. assessing the position of the zoid in relation to a socket) easier, faster, and more reliable, as mental rotation is now delegated to the artifact.

There are numerous kinds of epistemic actions[40] and the majority of them involve cognitive artifacts[41]. Some merely concern the interpretation of information, others the

[40] In the second chapter, I distinguished between two classes of cognitive techniques, those that substitute cognitive artifacts and those that concern interactive functional skills that allow us to exploit our environment for cognitive purposes. An epistemic action is an action whose function it is to improve cognition by extracting task-relevant information from the environment or by creating task-relevant

manipulation of artifacts, and yet others the offloading of information. Sometimes an agent needs to (quickly) obtain external information for whatever purpose. For example, when I need to know the time, a quick saccade to the clock on my desk suffices. But sometimes I need to perform a series of actions to obtain the information I need. For instance, when I need information about a particular philosophical concept, I get a particular book from my bookcase, open it at the index to look for keywords, then go to the specific page(s) on which those keywords can be found, followed by reading the relevant passages. In the first case, there was only one action involved and there was no direct manipulation of an artifact. In the second case, there was a series of nested actions involved and an artifact was deliberately manipulated in order to obtain the relevant information.

Other epistemic actions first concern the offloading of information onto the environment and then the intake of that information for some cognitive purpose. When I leave a rented DVD on my desk as a reminder to bring it back to the video store, I offload information onto the environment by creating a physical-spatial structure that contains task-relevant information. This is a rather simple epistemic action, as it involves only two subsequent actions: grabbing the DVD and leaving it on the desk. There are also more complex epistemic actions that concern the offloading of information. For example, when a designer sketches the rough outlines of a structure, then examines the ratios between elements in the design, re-sketches some of those elements, perceives the result, re-sketches, perceives the result, and so forth, until the correct structure has been sketched, numerous information-structures are offloaded. Creating this overall external information-structure involves numerous distinct but interrelated and nested actions.

Epistemic actions may thus concern the interpretation of information provided by cognitive artifacts, the manipulation of (cognitive) artifacts to obtain the relevant information, the manipulation of (cognitive) artifacts to create the relevant information, or a combination of these. In either of these cases, epistemic actions are

---

information in the environment. Epistemic actions are thus very similar to my notion of interactive functional skills.

[41] Epistemic actions may also involve cognitive naturefacts (e.g. navigating on the basis of celestial objects), human agents (e.g. asking a question), or involve a (re-)orientation of one's body towards something. These are epistemic actions because their goal is to change the input to the onboard cognitive system.

performed so that information can flow between agent and artifact. Epistemic actions are thus the building blocks for information flow in situated cognitive systems. In order to better understand information flow, we should conceptualize how agents interact with their artifacts. In Michaelian and Sutton's words: "The cognitive scientists' aim then is to examine the microprocesses of interaction across the diverse components of these distributed and heterogeneous cognitive systems, tracing for example the propagation and transformations of particular representational states across distinct (internal and external) media" (Michaelian & Sutton 2013, p. 5).

3.1.1 One-Way Information Flow

Below I identify four kinds of information flow, which should not be seen as clearly distinct, but as overlapping. They include one-way, two-way, and reciprocal information flow between agent and artifact, but also information flow in larger systems that comprise more than one agent and more than one artifact, which I call system information flow[42]. The first kind is characterized by a monocausal or one-way information flow from artifact to agent. Examples include clocks, compasses, road signs, maps, dictionaries, encyclopaedias, newspapers, websites, textbooks, documentaries, graphs, diagrams, manuals, timetables, et cetera. Humans make decisions and structure their actions on the basis of the information that such artifacts provide. We depart to the train station on the basis of a timetable and our watch, we take a left turn because the map says it is the shortest route to our destination, we assembly a piece of furniture in a certain way because the assembly instructions inform us to do so, and so forth. Such artifacts are usually open access systems in that they are open to anyone able to interpret the information. Further, and this is essential, the agent typically does not have any influence on the content and nature of the information. The information that such artifacts contain is fixed and not transformed during a task. Such artifacts and the information they provide are designed by other human agents including writers, journalists, designers, publishers, companies, governmental institutions, et cetera. One can thus argue that such cognitive artifacts

[42] The dimension of information flow is somewhat related to Sterelny's (2010) dimension of individual versus collective use, but is more fine-grained.

mediate information flow between the designer(s) of the information and its user(s)[43]. So, in this sense, the cognitive artifact is in between a designer and user.

In some cases, we interact bodily with the artifact in order to obtain the information we need. We interact bodily with compasses, rulers, maps, cookbooks, and manuals to get the information we need, in which case there are three steps involved: bodily interaction, perceptual intake, and action[44]. In other cases, we merely have to look at the artifact to extract the relevant information. Road signs, for example, need not be interacted with bodily to obtain the information we want as we just have to look at them, in which case there are two steps involved: perceptual intake and action[45]. However, note that not every deployment of a cognitive artifact results in an action. Occasionally, we are inhibited from performing an action. No-smoking signs and no-parking signs, for example, do not result in an action, but in an inhibition of an action (but, of course, only for those who would have otherwise smoked or parked). There might also be cases where there are neither actions nor inhibition of actions. A curious person who looks up the meaning or definition of a word in a dictionary, may do so just to satisfy her curiosity. Information thus sometimes has intrinsic value, i.e., it is valuable in itself and is not used for some other purpose. In most cases, however, information has extrinsic value: it is then used for completing a cognitive task.

3.1.2 Two-Way Information Flow

The second kind is characterized by a bicausal or two-way information flow, i.e., from agent to artifact and then from artifact to agent. Humans frequently offload information onto their environment to relieve their memory burdens, in that way creating cognitive artifacts such as post-it notes, notebook and diary entries, shopping-lists, to-do lists, and lists of addresses, birthdays, and telephone numbers. But also artifacts with system cognitive functions such as leaving car keys on a certain location

---

[43] There are cases in which the designer of the artifact also uses the artifact, for example, the writer(s) of a dictionary, encyclopaedia, or cookbook may use the books to look for information and to structure their actions in which case there is a two-way relation established, see next subsection.

[44] Although the information flow between agent and artifact is monocausal, there is strictly speaking not a monocausal relation with the artifact itself. Rather, a bicausal relation is established, because we have to causally and bodily interact with the artifact to obtain the information we need or want.

[45] Some theorists (e.g. Merleu-Ponty 2006) have argued that looking or perceiving is a form of bodily interaction.

or structuring drink glasses such that they correspond to the order of the drinks exhibit a two-way information flow structure. Artifacts in two-way relations are often tailored for individual use and are frequently not part of publicly available artifacts or representational systems (such as, e.g., road signs, clocks, and textbooks), although there are exceptions such as a shared diary. They are closed systems in the sense that the cognitive artifact is meant for an individual agent who has designed the informational content of the artifact for individual use. Once the information is offloaded, it remains fixed and is thus not transformed during a task. In one-way systems, designers outside the system have designed both the physical structure and informational content of the cognitive artifact. But in two-way systems, designers have designed the underlying physical structure of the artifact, e.g. the physical structure of a notebook or car keys. However, and this is essential, the informational content and its cognitive function is designed by the user, which is very distinctive from one-way systems[46].

In terms of repeatability, there are different versions of two-way relations. On one side of the spectrum there are one-offs like, for example, a post-it note with a brief reminder such as the date of a deadline. This example is a one-off, as there is only one cycle of offloading, intake, and action. On the other side of the spectrum there are often repeated interactions with a single artifact as in the case of Otto and his notebook. Otto writes important information in his notebook and then consults it to act on the basis of that information. He usually does not further manipulate the existing information in the notebook, but does on occasion add new information to it when he needs to do so. The content of new entries in the notebook usually does not depend on the content of previous entries. When Otto writes down the address of MoMa, it is because he knows that in the future he might be going to MoMa and for some external reason he is triggered to write down its address. It is most likely not because other information in the notebook triggered him to do so. In case of post-it notes, there is typically one cycle of offloading, intake, and action. But in case of Otto's notebook, there are various

[46] One can imagine cases in which another human agent, a secretary for example, makes a to-do list or diary entry for someone else, in which case there is no two-way information flow between agent and artifact. The artifact, then, mediates the information flow between two agents. Such cases therefore fall under one-way systems. But one can also imagine cases in which one-way systems are tailored to individual use, for example by highlighting a location on a map as a reminder of the location of a place of interest, or by making notes or additions in the sideline of a textbook. Such cases therefore fall under two-way systems.

distinct cycles of offloading, intake, and action, which are repeated over a certain period of time. However, it is important to note that the informational content of each cycle does not depend on the informational content of previous cycles. Hence, Otto and his notebook form a two-way system, just one that is often repeated. However, if Otto does use previous information in the notebook and further manipulates it, he and the information in his notebook then constitute a reciprocal system.

3.1.3 Reciprocal Information Flow

The third kind is based on a reciprocal information flow. Cognitive artifacts are sometimes integral parts of ongoing information-processing tasks. Writing an academic paper, making a PowerPoint presentation, solving a difficult calculation with pen, paper, and numerals, or designing an architectural blueprint of a building, often involves small incremental steps. We do not have a finished paper, presentation, calculation, or blueprint in our head and then fully offload it onto the artifact. Rather, we offload small bits of information onto the artifact, and the nature and content of the offloaded information contributes to and partly determines the next step in the overall process. For example, when writing an academic paper one often starts with a rough outline, which may prompt ideas about how to fill in the details. Filling in the details may then prompt an adjustment of the outline, which may in turn prompt further details. This process may continue for a number of cycles. Each step in the overall process builds and depends on previous steps. The human agent and cognitive artifact continuously exchange information and so there is a reciprocal and cumulative information flow that constantly transforms the situated cognitive system. There is, in Clark's words, "continuous reciprocal causation" between agent and artifact (Clark 1997, p. 164). So, contrary to one-way and two-way systems, the external information in reciprocal systems is continuously transformed during a task.

Like information flow in two-way systems, reciprocal information flow often takes place in a closed system in the sense that the cognitive artifact is meant for a single agent who has designed the informational content of the artifact for individual use. In two-way relations, there were three steps involved: offloading, intake, and action. This is roughly the same for reciprocal relations, except that each cycle depends on the outcome of the previous one. The cycles are thus *interdependent*. For this reason, the functional and informational integration between agent and artifact is significantly closer than in one-way or two-way systems. It is not a mere exchange of information

between two entities, as in two-way systems. What is offloaded onto the artifact in a given cycle depends on what is offloaded in the previous cycle(s) and, therefore, the degree of integration is considerably higher. In fact, this integration is so dense that it is better to conceive of agent and artifact as one cognitive and information-processing system.

Finally, the outcome or result of one-way and two-way systems is often an action or an inhibition of an action. However, the outcome of reciprocal systems is a cognitive artifact in itself (e.g., a paper, PowerPoint presentation, answer to a calculation, or architectural blueprint) and not an action. The developed cognitive artifact may be used to inform human action, both for the designer of the artifact and for others. For example, in the case of a PowerPoint presentation, it serves as a memory aid for the presenter, which then becomes a two-way system, but also to convey information to an audience, for which is a one-way system.

3.1.4 System Information Flow

The above three kinds of information flow concern situated cognitive systems that are comprised of an embodied agent and a (set of) artifact(s). However, information quite often flows in systems that are comprised of more than one agent and more than one cognitive artifact. These are the cases in which Hutchins (1995a, 1995b), Nersessian (2005, 2006, see also Nersessian et al 2003), Evelyn Tribble (2005, 2011), and others are primarily interested in. Examples include a team of navigators trying to navigate a ship through a harbor, pilots in the cockpit of an airplane, researchers in a scientific laboratory, or actors interacting with theatre architecture and other artifacts. These are cases of distributed cognition in which there is a collective of agents that tries to solve a particular problem or perform a certain cognitive task by using a variety of cognitive artifacts. Within those distributed systems, information flows in many directions (sometimes simultaneously) and is often transformed and reformatted (see next section).

It could be argued that within those larger distributed cognitive systems, there are often agents interacting with single artifacts, for example a navigator using an alidade to measure the bearings relative to the ship's head, a pilot reading a radar, a scientist using a flow loop to generate representations of blood cells, or an actor interacting with a playbook. Agent-artifact systems are thus often the building blocks of larger

distributed cognitive systems, implying that a better understanding of agent-artifact systems can have a trickledown effect on better understanding larger distributed systems. Furthermore, the way agents interact with those artifacts is influenced by the history and social organization of the larger system, which should be taken into account when analysing single agent-artifact subsystems within larger systems.

3.1.5 Information Trajectories and Transformations

Units of information (either representational or ecological) are propagated within situated cognitive systems, i.e., they have particular pathways or routes. These may be referred to as *information trajectories* (see also Hutchins & Klausen 1996). In one-way systems, this trajectory is fairly short, as information is propagated from artifact to agent. In two-way systems, this trajectory is twice as long, as information is propagated from agent to artifact to agent. In reciprocal systems, this trajectory can be relatively short (e.g. when performing a short calculation with pen, paper, and numerals), but it can also be rather long (e.g. when sketching, the amount of interdependent cycles of informational offloading and intake can be quite substantial). In larger distributed systems, information can also have short or long trajectories. A pilot might, for example, briefly look at the radar to check if there are no changes in the weather situation, in which case a one-way information trajectory is established. But a pilot might also look at the radar, discuss the readings with the co-pilot, make some adjustments to the radar's settings because the task-relevant weather information is not adequately visible, and then make some notes in the plane's log on the basis of the discussion with the co-pilot and the readings of the radar. In this case, the trajectory is significantly longer and involves more informational nodes.

Moreover, information can transform and be reformatted at certain informational nodes in the system. This may concern transformation of representational format or of informational content. Pierce (1935a, 1935b) distinguished between three kinds of representations: icons, indices, and symbols. Icons are isomorphic to their target (e.g. maps), indices have a direct causal connection to their target (e.g. weather vanes), and symbols obtain their meaning through logical rules and agreement (e.g. language or mathematics). When trying to convert a table into a graph with the aid of a spreadsheet program, the table is transformed into a different representational format, i.e., from a set of numerals (i.e. symbols) into a graph (i.e. an icon). In this case, informational

content remains the same, but its representational format is transformed. But one can also transform informational content, e.g. when changing the data in the table.

The above example concerns transformation within an agent-artifact system, but there are also more complex forms of transformation within larger distributed systems. Consider Nersessian's (2005) example from a biomedical engineering laboratory. A device called the "flow loop" emulates the shear stresses experienced by cells within blood vessels. During a simulation, the flow loop manipulates particular biological constructs which are representations of artery walls. After the simulation, the constructs are examined with instruments such as a confocal microscope, which generates high-definition images of the constructs. This simulation process allows the researchers to gain information about, e.g., the number of endothelial cells and about the direction of filaments in relation to the blood flow. Thus, Nersessian concludes, "the *representations generated* by the flow loop *manipulations* of the constructs are *propagated* within the cognitive system" (Nersessian 2005, p. 43, original italics). In this example, there is a fairly long information trajectory and there are different kinds of transformation at work.

3.2 Speed of Information Flow

This dimension concerns how fast information flows between the elements in a situated cognitive system. How quickly a cognitive task is performed is often important for realizing a cognitive purpose. When playing Tetris, for example, a user only has a few seconds to decide in which orientation the zoid should be dropped into a lower template (Kirsh & Maglio 1994). Likewise, pilots landing an airplane need to perform various cognitive tasks within a certain time frame (Hutchins 1995b). The same is true for fast rally drivers, scientists performing experiments, engineers designing objects or structures, and many other tasks. These examples show that there is time pressure for many cognitive tasks.

Information flows between the components in a situated cognitive system and how fast this information flows depends on both the informational properties of the artifact and the cognitive profile of its user. Some people read quickly, while others do not. Some people interpret a map in one glance, while others have to study it before they know where to navigate (Skagerlund, Kirsh & Dahlbäck 2012). Humans have thus different interpretative and interactive skills, which partly determine how fast information is

taken onboard and processed. The degree of informational transparency is also relevant here (see section 3.8). Some information is easier to interpret than other. The higher the informational transparency, the easier the information is to interpret, and the higher the speed of information flow. So speed of information flow depends, on the one hand, on the cognitive and interpretation skills of the human agent and, on the other hand, on the informational and representational nature of the cognitive artifact (see also section 3.8). But contextual factors such as background noise may also influence speed of information flow, since one's concentration and thus also one's ability to interpret information is influenced by it.

Conversely, the speed with which one offloads information onto an artifact is also important. Again, this depends on properties of both the agent and artifact. Certain devices have input methods that are more efficient than others. A desktop computer has a keyboard that is geared towards quick data input, a tablet has a virtual keyboard that is much less efficient, and a smartphone has a virtual keyboard as well, but one that is much smaller and thus significantly less efficient. Some computing devices have auditory input methods which are potentially much quicker than conventional methods, because most people can speak quicker than they can type or write. But equally relevant are the interactive skills of the agent. Some people write or type considerably quicker than others, which often depends on how one's body schema and motor programs have developed.

3.3 Reliability

This dimension concerns how reliable an artifact is available for a user in a given context. Reliable access to external information is important for understanding how and how often an epistemic (inter)action unfolds. "It happens that most reliable coupling takes place within the brain, but there can easily be reliable coupling with the environment as well" (Clark & Chalmers 1998, p. 11). In the literature on extended cognition, the dimension of reliability is often discussed in terms of how consistently and reliably an agent has access to a cognitive artifact (Clark and Chalmers 1998; Wilson and Clark 2009). But there are several other features relevant to reliable information access that are often not discussed in the literature.

First, the cognitive profile of the human agent partly determines the necessity for information access. For certain cognitive tasks we highly depend on cognitive artifacts.

Finding your way in an unfamiliar city without the support of road signs, maps, and/or navigation systems; multiplying 3567 x 5674 without the help of pen and paper or a calculator; or remembering all your appointments for the entire month is for most people impossible. We highly depend on cognitive artifacts to help us perform those cognitive tasks and without such devices our cognitive capacities would just not be the same. The degree of dependency partly depends on one's cognitive profile. Some people, like Otto, have bad memory capacities and therefore rely and depend on memory aids such as notebooks, post-it notes, diaries, other people, and other reminders. Other people have bad mathematical skills and rely and depend more on calculators or perform calculations with pen and paper. While yet other people have bad navigation skills and rely and depend on navigation aids such as road signs, maps and navigation systems. There are also people who have better memory, mathematical, and navigational skills and do not or rely less on external artifacts.

Second, reliability depends on the kind and properties of the artifact. Due to their physical properties, some artifacts provide better information access than others. Take diaries, for instance. As long as one does not forget to bring one's analogue diary when needed, it provides reliable access to the information in it. In contrast, digital diaries embedded in one's smart phone, tablet, or other electronic device, in one sense, provide less reliable access, because they are inaccessible without electricity. So one not only needs to remember to bring the device when needed, but also to charge it when the battery is empty. Further, digital cognitive artifacts can potentially malfunction in more ways than analogue ones. So next to battery issues, there may be numerous software and hardware issues that prevent one from accessing one's digital diary. Software and hardware issues are irrelevant for analogue diaries. But, conversely, digital diaries such as Google Calendar are online systems that store information in the cloud and are therefore less susceptible for theft, loss, or damage than analogue diaries. Even if one loses one's wearable computing device or if it gets stolen, the information is still available in the cloud. Analogue diaries lack these properties.

The degree of portability is also an important physical property for information access. Cognitive artifacts like smart phones, slide rulers, compasses, or watches are worn or carried on one's body and are thus very portable. As a result, they are (when fully operative) easily, repeatedly, and reliably accessible. The physical design of these artifacts is such that they are small, light, and (ideally) ergonomic, because the

epistemic actions that are performed with these artifacts require them to be small, light, and ergonomic. If not, then they are not suitable for their function. A large and heavy compass, for example, is, in virtue of its physical properties, not particularly portable and thus non-functional for hiking. One could imagine a fully functioning fifty kilogram compass that would be able to fulfil its representational function, but due to its physical properties, unable to be of any help for bushwalkers and so in practice non-functional for the cognitive purpose of measuring one's position in relation to the north pole. Tablets, laptops, analogue diaries, and books are still rather portable but often not (though in some cases they are) worn or carried on one's body and thus a bit less (often) accessible. Desktop computers, DNA-sequencers, and radio-telescopes are non-portable and only accessible at their fixed location[47]. The point of these examples is that partly due to their physical properties (i.e. weight, size, design), cognitive artifacts are able to fulfil their function.

Third, the context and kind of epistemic action are relevant for reliable information access. A carpenter only brings his ruler when he effectively needs it, which is during work. Carpenters only need access to rulers when they need to perform the epistemic action of measuring the length of some object. Such epistemic actions are frequently performed during work and thus in a work-environment. Carpenters presumably do not bring rulers to the supermarket or dinner parties, because there is nothing for them to measure. Likewise, a bushwalker only needs to bring a compass when she is bushwalking. Bringing a compass on a city trip seems a bit redundant. And a biochemist only needs access to a DNA-sequencer when she is performing research that for some goal requires the sequence of a certain piece of DNA. So necessity of information access partly depends on the kind of epistemic action and context. Certain epistemic actions are therefore only performed in particular contexts.

3.4 Durability

There are two sides to durability. First, the durability of the artifact itself. Second, the durability of the relationship with the artifact. Certain cognitive artifacts are highly durable, while others are less so. When handled carefully, textbooks, abacuses, and rulers can potentially last for decades, whereas analogue diaries last for roughly a year,

[47] Such devices are of course portable in the sense that they, like most artifacts, can be transported, but they are informationally inaccessible while being transported.

and shopping-lists and to-do lists often last for just a few hours. This depends on both the material quality and properties of the artifact as well as the purpose of the epistemic action. Generally, the more robust the material quality of the artifact, the more durable it is. When the artifact is not (sufficiently) durable, it may result in malfunction (Preston 2009).

But, more importantly, the durability and repeatability of our relationship with cognitive artifacts often depends on the kind of epistemic action (and its epistemic purpose) one performs with it. A shopping-list does not need to be very durable, because after having bought the needed items, it has fulfilled its epistemic purpose. A desktop computer, in contrast, *does* need to be durable because we need it for many kinds of epistemic actions for a long period of time. Wilson and Clark (2009) introduce a trichotomy between one-offs, repeated, and permanent relationships to cognitive artifacts. Shopping-lists are typically one-offs. Abacuses or compasses, however, are frequently re-used because they are devices that are utilized many times for the purpose of calculating or navigating. But some cognitive artifacts enter into permanent and highly durable relationships with their users. Otto and his notebook, a carpenter and his ruler, and an accountant and her calculator enter into long-lasting and interdependent relationships.

A further distinction can be made between the duration of the temporary coupling of a specific agent-artifact system and the repeatability of the same system. We frequently use calculators to help us solve difficult calculations, which is often a short process. As soon as we have the answer, the system is decoupled. However, throughout our lives we repeatedly and consistently use calculators, so there are many cycles of coupling, decoupling, and recoupling with certain cognitive artifacts distributed over long periods of time. Conversely, there are also cognitive artifacts with which we couple only once or perhaps only a few times. A compass, for instance, is an artifact most people have probably used once or twice. Only bush walkers, scuba divers, and sailors use it quite often. Most people do not need a compass to get around in the world and so there are situated cognitive systems that are for reasons of necessity not repeated. The duration of coupling also depends on task and artifact. Using a calculator may only take a few seconds, whereas writing a PhD thesis on a desktop computer may take many years.

3.5 Trust

The notion of trust in the literature on extended cognition concerns our attitude towards the truth-value of information. When we trust information, we typically think it is true. When we distrust information, we either think it is false or we are not sure whether it is true. There are at least two ways in which we can come to (dis)trust information. First, by consciously evaluating it and then come to the conclusion that the information is either trustworthy or not trustworthy. Second, by assuming it is either trustworthy or not trustworthy without consciously evaluating it. The first type of trust may be called explicit (dis)trust and the second may be called implicit (dis)trust.

Some information we trust implicitly because we have endorsed it somewhere in the past and wrote it down because of this endorsement. This is true for Otto's notebook, diary entries, shopping-lists, and the information on post-it notes, as we usually (though not necessarily) do not consciously evaluate information we wrote down ourselves. Other information we trust implicitly because many people rely on it for their actions. This is true for timetables of trains, dictionaries, encyclopaedias, and maps, which are used and shared by many humans. Because these artifacts are shared with many others, and many people rely on them for their actions, there is often no reason to think that the information they contain is false or incorrect (Sterelny 2010, contrast Sterelny 2004). But there are certainly exceptions: Wikipedia, for instance, is used and shared by many people, but given its great variability and constant change, it is in some cases still not a trustworthy source of information (Magnus 2009; Simon 2013).

In two-way and reciprocal situated cognitive systems, we implicitly trust the information because we have endorsed it in the past and because we offloaded it ourselves, but we also implicitly trust it because we believe the information is private and has not been tampered with. Consider a brief example: in Australia there is a TV commercial for smart phones in which a parent goes shopping with a shopping-list composed on a smart phone. The application is connected in real-time to the desktop at home where his son deliberately changes the digital shopping-list to include items he desires. This example shows that with new digital cognitive artifacts exhibiting networking abilities such as smart phones and tablets, informational privacy and security become increasingly important for trust in information (Floridi 2005, 2006).

Privacy and security issues are less likely to emerge when using analogue shopping-lists, which are identifiable by means of one's handwriting (Parsell 2007). So the nature and properties of the artifact partly determine how relevant informational privacy and security are for establishing a trust relation with the artifact and the information it carries.

3.6 Procedural Transparency

There are two types of transparency that are relevant for cognitive artifacts, namely, procedural and informational transparency. In chapter 2, it was shown that embodied tools such as, e.g., screwdrivers, spoons, cricket bats, and hammers, transform the body schema. For developmental reasons, body schemas are flexible as to incorporate tools into the subpersonal representation of the body and its capabilities for action. Those tools, then, are not experienced as external objects with which one interacts, but one interacts with the environment through those tools. When a tool is incorporated into the body schema of its user, it becomes transparent-in-use. When that happens, we no longer need to consciously think about how to interact with the tool, interaction goes smoothly and the tool withdraws from attention, but it does not become invisible (see also Ward & Stapleton 2011).

A similar phenomenon occurs with cognitive artifacts which I will call "procedural transparency". Procedural transparency (see also Clark 2008b; Sutton 2009a) concerns the effortlessness and lack of conscious attention with which an agent deploys a cognitive artifact. Cognitive artifacts that are procedurally transparent are not incorporated into the body schema and are not experienced as an extension of the body and its motor system, but are nonetheless transparent-in-use. Otto, for example, is so adapted and familiar to using his notebook that he will consult it automatically when he needs to do so. His perceptual-motor processes are proceduralized to such an extent that he does not consciously think about how to retrieve information from his notebook. So the retrieval process is not a two-step process in which Otto first believes that the address of MoMa is in his notebook and then looks up and interprets the information to form his second belief, namely, that MoMa is at 53rd street. It is a proceduralized and transparent process. In Clark's words: "the notebook has become transparent equipment for Otto, just as biological memory is for Inga" (Clark 2008b, p. 80).

Likewise, maps, textbooks, compasses, abacuses, computer applications and countless other cognitive artifacts, typically do not need conscious attention in order to effectively use them and can thus be said to be procedurally transparent. It must be noted, though, that achieving procedural transparency, in most cases, needs training and takes a considerable amount of time. Novice cognitive artifact-users typically lack procedural transparency. The first time I pick up a compass or abacus, it is difficult for me to effectively use it, because there is no procedural transparency. But as my compass and abacus-using skills develop, I become more fluent in successfully using the artifact. Thus the more a user becomes an expert at using a cognitive artifact, the more fluent and efficient the cognitive task is performed.

3.7 Informational Transparency

Informational transparency concerns the effortlessness with which an agent can interpret and understand external information (see also Cheng 2009). In the previous chapter, I developed a taxonomy in which I first distinguished between representational and ecological cognitive artifacts and then made further distinctions within those two genera. In case of representational artifacts, I distinguished between icons, indices, and symbols. And in case of ecological artifacts, I distinguished between structural and spatial artifacts. Informational transparency typically occurs with all these kinds of cognitive artifacts. Icons are defined by exhibiting isomorphism with their target. Generally, but not necessarily, the higher the isomorphism between icon and target, the higher the informational transparency. Photographs, for example, display a high degree of isomorphism with their target and are therefore highly transparent, whereas an fMRI-scan is also isomorphic to its target but needs more interpretative skills and training to meaningfully interpret (see for skilfully interpreting fMRI scans, Alač & Hutchins 2004). For this reason, it is less transparent, at least for people who lack those interpretative skills. Both the skills of the agent and properties of the icon determine its informational transparency. Thus the easier it is for an agent to understand the salient properties of an icon, the more transparent it is.

Indices have a direct causal connection between index and target. The degree of informational transparency depends on an understanding of this causal connection, i.e., the better the causal connection between target and index is understood by an agent, the more transparency the index is. So, for example, the causal connection between the state of a weathervane and the direction of the wind is for most people quite easy to

understand and therefore quite transparent. The causal connection between an electrical current and the reading on a voltmeter is probably a bit less easy to understand, as electricity is a more complicated phenomenon than wind direction. And the causal connection between a gel electrophoresis pattern and protein structure and protein size is, unless you are a molecular biologist, quite difficult to understand and so it is close to opaque. Thus, again, one's interpretative skills and informational properties of the artifact determine the degree of transparency.

In case of symbols such as language and mathematics, the degree of informational transparency depends on the degree of understanding of the rules and conventions and the amount of shared use that determine the meaning of symbols. For a native English speaker, for example, English is highly transparent because the rules and conventions that determine the meaning of language (i.e., syntax and semantics) are fully understood and, furthermore, language is shared with a community of language-users. For a non-English speaking person, English is opaque as the rules are not understood and the language is not shared. The same applies to mathematics, so for someone who has not been trained in mathematics, mathematical symbols and numbers are opaque. Finally, as pointed out in the previous chapter, a token representation may display a combination of iconic, indexical, and symbolic properties. As a result, the informational transparency of some representations may depend on a mixture of isomorphism, an understanding of the causal connection between vehicle and target, and an understanding of the rules and conventions that determine the meaning of some representations.

Ecological artifacts, too, are informationally transparent. However, given that such artifacts do not contain representations but exhibit ecological information, their transparency does not concern transparency between vehicle and target. Spatial ecological artifacts obtain their cognitive function in virtue of their physical-spatial structure. This physical-spatial structure is usually made by the user of that structure. So, for example, when putting the pieces of a dismantled bicycle on a sheet of newspaper in such a way that the location of those pieces indicate the order in which have to be put together, an agent improvises a physical-spatial structure which aid in performing a cognitive task. Because such structures are made by the users themselves and therefore geared towards efficient use, it is clear what those structures are meant

to do. So for the user, the information provided by the ecological artifact is quite transparent.

3.8 Individualization

This dimension concerns how much an agent has personalised or individualized an artifact as to make it better suitable for the agent's idiosyncratic cognitive goals. Sterelny (2010) presents a spectrum in which some cognitive artifacts are interchangeable, others are individualized, and yet others are entrenched. For Sterelny, individualization is changing, adjusting, or fine-tuning the informational properties of the artifact such that its use is more effective and efficient for realizing its cognitive function. He argues that most of the books in his professional library are interchangeable, but some of them are massively individualized with underlining, highlighting, comments, and post-it notes. These adjustments essentially make sense to Sterelny and are less useful and valuable to others. Similarly, Otto's notebook is highly individualized and is useful only for Otto, although others may still be able to read the notebook, only Otto uses it to aid his memory and structure his actions. My tablet computer is fairly individualized: it has applications that I have downloaded and installed to fit my specific needs such as the weather forecast and train timetables for Sydney, and specific websites, documents, and books. But although it is individualized, most applications are still easily usable by others. In contrast, road signs, speed dials, weathervanes, newspapers, and textbooks are not individualized (and thus interchangeable) and accessible for most people. Like transparency, individualization of cognitive artifacts often takes a certain period of time, and highly individualized cognitive artifacts are in close equilibrium with the cognitive profile of their user.

Entrenchment of cognitive artifacts implies a close equilibrium between agent and artifact in which *both* have been transformed in order to ensure the best possible fit between agent and artifact. Sterelny acknowledges that his individualized books are not entrenched in the sense that his professional routines and habits have not been adjusted to those books in the same way as those books have been adjusted to Sterelny. So, he has individualized his books, but his books have not individualized him, or at least not sufficiently. But, according to Sterelny, there may still be cases of entrenchment concerning books. For a Locke scholar, Locke's *oeuvre* may have transformed the routines of the scholar in the same way as he or she has transformed Locke's *oeuvre* in the sense of highly individualizing his works by underlining,

highlighting, comments, and so on. A more obvious and clear example of an entrenched cognitive artifact is Otto's notebook. The information in his notebook is only meant for Otto himself and is specifically geared to his needs and desires, so it is highly individualized, and his behavioural and cognitive routines are sculpted by his notebook, so it is entrenched as well. Generally, interchangeable artifacts concern one-way, two-way, or reciprocal information flow, whereas individualized and entrenched artifacts concern either two-way or reciprocal information flow.

3.9 Transformation

This dimension concerns how interacting with artifacts transforms the representational properties of the brain. We have seen that body schemas are fluent and flexible so that tools can be incorporated into the subpersonal representation of the body and its capabilities for action. Tool-use thus transforms the body schema. Likewise, the use of cognitive artifacts transform the representational and cognitive capacities of the human brain. Clark (1997), Stanislaw Dehaene et al (1999), Helen de Cruz (2008), Menary (2010c), and Michael Kirchhoff (2011), amongst others, have argued that external representational systems transform the brain's capacities. During ontogenetic development, we interact with public representational systems such as language and mathematics. By so doing, we soak up and learn to think in those systems and the brain takes on the representational properties of those systems.

Language and mathematics are examples of external representational systems with which we interact substantially for a long period of time, both phylogenetically and ontogenetically. In ontogeny we call this period education. A considerable amount of research has been done on the transformation effect of those systems on our brain and cognition (e.g., Dehaene et al 1999). Other cognitive artifacts and representational systems such as road signs, maps, graphs, diagrams, computers, and design programs have presumably also a transformation effect on our representational and cognitive capacities. For example, after navigating a city with a map for a certain period of time, the interaction with the map and the city has changed our internal spatial representation, i.e., our cognitive map, of parts of the city. At a certain point, we no longer need the actual map to navigate and we have to a certain degree internalized the information of the map. Likewise, interacting with computers for many hours a day almost certainly transforms our neuronal structures and cognitive capacities (see also Carr 2011). Engineers, for example, spend many hours a day designing objects and structures with design programs. It is not unlikely that after a certain period of training

and practice their brains take on the representational properties of the software program, i.e., engineers learn to think in terms of the representational systems of the program (see, e.g., Eastman 2001). Such transformations seem to be a consequence of long-term interaction with cognitive artifacts over ontogenetic time. The history of interaction is thus highly important for understanding the current cognitive capacities of an embodied agent.

As I have pointed out in section 3.7, interacting with cognitive artifacts not just transforms the brain's representational properties, it also transforms our embodied interactive skills. An obvious example is a personal computer. An agent who has never used a personal computer will have trouble interacting with it. The mouse, keyboard, and screen will be difficult to use and it takes a certain amount of time to learn to (fluently) use a computer. Our body schema and motor programs need to adapt and transform in order to be able to successfully interact with computers. Likewise, as outlined in chapter 2, compasses, rulers, navigation systems, voltmeters, and countless other cognitive artifacts require for their use certain embodied interactive skills that are developed and transformed during the use of such artifacts.

It is, however, not only the human component of the situated system that transforms its representational properties and capacities. The artifactual component transforms its representational properties, too, i.e., cognitive artifacts are often not static and fixed but active and dynamic. The representational properties of post-it notes, slide rulers, and textbooks, for instance, are fairly stable and fixed, but smart phones, tablets, laptops, and other computing devices are very dynamic in their representational properties. We can transform and adjust their representational properties to our own needs and desires, and it is frequently *because* we act on those artifacts and the information they contain that they have dynamic and changing representations. Similarly, informational transformation also occurs when using ecological cognitive artifacts. When consistently leaving car keys on a particular spot in one's apartment, the informational properties of that ecological information-structure are quite static. But when playing Tetris, the zoids constantly change position and orientation due to the agency of the user, i.e., by pushing a button to rotate the zoid. So when talking about the dimension of transformation, one can make a distinction between neural, skill, and artifactual transformation, demonstrating that most situated cognitive systems are quite dynamic.

To conclude this section, I have now outlined ten dimensions for conceptualizing the integration between embodied agents and cognitive artifacts. In the following section, I briefly examine the relations between some of these dimensions and argue that integration should be seen as an inherently multidimensional phenomenon. In the next chapter, these dimensions are used to conceptualize integration in four case studies.

4. Relations between Dimensions

In discussing a preliminary version of the multidimensional framework developed in this chapter (see Heersmink 2012b), Robert Clowes (2013) claims that the proposed dimensions are relatively independent and that more practical work and conceptual analysis is needed to clarify which dimensions are more important. Clowes himself suggests that transparency and trust are, at least in cases of external memory, of central importance. Although he does not distinguish between procedural and informational transparency, I take it he is referring to procedural transparency, as he illustrates his point by making an analogy to Heidegger's notion of a hammer being ready-to-hand. My view is that it is problematic to prioritize certain dimensions because it encourages us to overlook other dimensions. If we would focus only on trust and procedural transparency, we would miss out on how information trajectories are established, how artifacts are individualized, how they transform our onboard cognitive system, and so forth. All these things are essential for better understanding integration. Thus, what we need is a picture of integration not in terms of necessary and sufficient conditions (see also Sutton et al 2010), but as an inherently multidimensional phenomenon without prioritizing any of these dimensions or group of dimensions.

For analytical purposes, I discussed each dimension separately, but some overlap and interact. I now look at some of these interactions and point out how the dimensions are related. The dimensions of information flow and speed of information flow are concerned with information trajectories, i.e., with how, how fast, or how much information flows in a situated system. Information trajectories themselves supervene on how the overall system is structurally and functionally organized. In other words, how information trajectories are established depends on the material and functional constitution of the integrated situated system.

Reliability and durability are both concerned with information access. If an artifact is not easily and reliably accessible, then it is hard to establish a durable relation to it. Further, reliability and durability often result in individualization. The more often a certain cognitive artifact is used, the more likely it is that it will be individualized and perhaps in some cases even entrenched. But this need not be the case. There are often-used cognitive artifacts that are neither individualized nor entrenched such as clocks and speed dials. Individualization and entrenchment frequently result in cognitive transformation. Again, the more often we use a certain cognitive artifact, the more likely it is that the human brain soaks-up the representational properties of the artifact. This happens with language and mathematics, but arguably also with maps, abacuses, design programs, and perhaps with graphs, pie charts, diagrams and other illustrations as well. Individualization frequently results in both trust and procedural and informational transparency. Individualized cognitive artifacts, including diaries and notebooks, are designed by the user of the artifact and thus almost automatically trusted and transparent in use, as well as transparent in interpretation. We do not need to think about how to use such artifacts, and the information they carry is trusted because we wrote it down ourselves.

How the overall system is structurally and functionally organized is interwoven with transformation. In other words, when components of situated systems are transformed (either embodied brains, skills, or artifacts), the structural and functional organization are typically transformed as well. Thus, when one's brain structures transform due to artifact-use, when new skills or interactive cognitive techniques are learned, or when the artifact's informational properties are transformed, then the overall structural and functional organisation is usually transformed, too. For example, when a novice computer-user is over time transformed into an expert computer-user, the novice's brain structures and embodied skills are changed such that the computer almost becomes a different cognitive artifact. The embodied skills one has partly determine the affordances of a cognitive artifact. A novice cannot perform the same cognitive tasks with a computer as compared to an expert, due to a lack of skill and thus also a lack of affordances. How the overall system is organised and transformed over time also relates to individualization. Typically, when agents individualize an artifact, they change its functional and informational properties so that it is better equipped to complement an agent's own properties, i.e., to ensure a better fit between the cognitive profile of the agent and the informational and functional properties of the artifact. And, finally,

informational transparency often results in a higher speed of information flow. The idea being that the easier information is to interpret and understand, the faster we can take it onboard and process it. Although there are more interactions between the dimensions, these are the most obvious ones.

Furthermore, situated cognitive systems can shift from one region of this multidimensional space to another. When a particular artifact is used for a longer period of time and it becomes gradually more efficient, durable, individualized, transparent, and trustworthy, the system becomes increasingly more integrated. As a result, the situated system will shift to a higher region in the multidimensional space. Highly individualized cognitive artifacts are likely to maintain a stable relation with its user and, consequently, populate a given region in the space for a long period of time, but most situated systems are frequently shifting from one region to another. This is so because a large part of these systems are quite dynamic in nature, constantly changing their functional and informational properties, and renegotiating existing functional and informational equilibriums. But although most of these systems are quite dynamic, I think that, ontogenetically, there is a tendency for situated systems to shift from lower to higher regions of this space. Otherwise put, it is much more common for systems to shift from lower to higher regions than vice versa.

5. Concluding Summary

This chapter began by presenting complementarity-based extended cognition theory. On a complementarity view, biological and artifactual elements in situated cognitive systems need not exhibit similar properties or functionalities. Rather, embodied agents deploy the different functional and informational properties of cognitive artifacts to complement their onboard cognitive capacities. Such a view encourages the study of the variety of situated cognitive systems by focussing on the particular properties of both agents and artifacts and by focussing on the functional integration of the two.

To better understand and conceptualize complementarity and functional integration, this chapter developed a multidimensional framework. This framework consists of the following dimensions: epistemic action and information flow, speed of information flow, reliability, durability, trust, procedural transparency, informational transparency, individualization, and transformation. All these dimensions are matters of degree and

relational in the sense that they emerge out of a specific cognitive interaction between agent and artifact performed in a particular context and with a specific cognitive purpose in mind. Importantly, these dimensions are not meant as necessary conditions for cognitive extension and thus do not provide a clear set of conditions to distinguish between cases of embedded and extended cognition. On my view, situated cognitive systems merely populate a certain region in this multidimensional space. The higher a system scores on these dimensions, the more tightly coupled the system is and the deeper the artifact and its user are integrated into a larger cognitive system. To illustrate how this framework can be put to work, in the next chapter four case studies are performed focussing on cognitive artifacts in scientific practice.

6

Cognitive Artifacts and Complementary Integration in Scientific Practice

1. Introduction

In this thesis, I began by reviewing several characterizations and classifications of components of situated cognitive systems. I critically evaluated current categorizations of cognitive artifacts and then developed a preliminary taxonomy that distinguished between artificial (i.e. artifacts and techniques), natural, and social scaffolds (chapter 2). Next, I further developed one category of this preliminary taxonomy by focussing on cognitive artifacts. I first defined the category of cognitive artifacts as a functional kind, one that includes artifacts with both proper (selected) and system (improvised) functions (chapter 3). Thereafter, by drawing on artifact categorization in archaeology, I developed a detailed and systematic taxonomy of cognitive artifacts. In developing this taxonomy, an information-centered approach was taken: I took as my point of departure the specific informational properties of cognitive artifacts and then taxonomized them on the basis of those properties, not the properties or goals of the agents that design, select, improvise, or use them (chapter 4). Having developed this taxonomy, I then briefly compared parity-based and complementarity-based extended cognition theory. I critically evaluated the parity principle and argued in favour of complementarity-based extended cognition theory. To better understand and conceptualize complementarity and integration between agent and artifact, I developed a multidimensional framework. This framework consisted of the following dimensions: epistemic action and information flow, speed of information flow, reliability, durability,

trust, procedural transparency, informational transparency, individualization, and transformation. These dimensions are the tools to conceptualize the degree of integration and jointly constitute a multidimensional space in which situated cognitive systems can be located (chapter 5).

Up to this point, this thesis has used examples of cognitive artifacts in a variety of settings and provided a general account of such artifacts and their relation to human agents. In this chapter, the focus will shift to the use of cognitive artifacts in scientific practice, particularly molecular biology laboratories (MBLs)[48]. One of the reasons for focussing on laboratories is because they are paradigm cases of environments in which a large variety of cognitive artifacts are used. Laboratories therefore provide an excellent context for taxonomizing cognitive artifacts and conceptualizing integration between agents and artifacts. Further, the cognitive processes of scientists when performing research are not different in kind than those in ordinary contexts, but can be seen to be part of a continuum (Nersessian 2009). Likewise, the cognitive artifacts used in laboratories to aid and complement those cognitive processes are also not different in kind than those used in ordinary contexts. They may exhibit much more complicated physical structures, but they can be classified in terms of the taxonomy. A second reason for focussing on laboratories is because a relatively small body of literature exists on (a) classifications of scientific instruments (Baird 2003, Harré 2003; Nersessian 2005) and (b) distributed cognition in scientific practice (Giere 2002a, 2002b; Magnus 2007; Nersessian 2005, 2009). I critically engage with this body of literature and compare it with my taxonomy and framework for conceptualizing complementary integration of agents and artifacts.

I proceed as follows. The first step in my analysis is to provide some structure and systematicity to the large variety of cognitive artifacts used in MBLs. In order to this, I begin with reviewing a number of current classifications of scientific instruments and point out that they overlook some categories of cognitive artifacts (section 2). Next, I

[48] The notion of a MBL should be seen as broadly construed. It may include research on genetics, biochemistry, cell biology, biomedical engineering, and other related fields. I use a broad notion of molecular biology as the cognitive artifacts used in different kinds of biology-related research are very similar. Perhaps even research in certain chemistry and physics labs make use of similar cognitive artifacts.

apply the taxonomy to classify some of the cognitive artifacts in MBLs[49] (section 3). Having done this, I then review how a number of theorists have used a distributed cognition approach to conceptualizing certain aspects of scientific practice. I critically evaluate these approaches and argue that they lack the tools for conceptualizing when a cognitive artifact is a proper part of a distributed cognitive system as opposed to it being a mere cognitive scaffold (section 4). The last step in my analysis is to conceptualize the degree to which cognitive artifacts in MBLs and their users are integrated into extended/distributed cognitive systems by applying the multidimensional framework to four case studies: (1) virtual models of protein folding, (2) pH-meters, (3) laboratory notebooks, and (4) organised workplaces (section 5). This chapter ends with giving a concluding summary (section 6).

2. Kinds of Scientific Instruments

With the exception of New Experimentalism (e.g. Hacking 1983; Ackerman 1985), the role of instruments in scientific practice has not played a prominent role in traditional philosophy of science (Boon 2009; compare Gooding, Pinch & Schaffer 1989; Ihde 1991; Pitt 2000, 2011). Only recently have philosophers of science became interested in instruments and their functional role in experiments and the creation of knowledge (see e.g. Radder 2003; Baird 2004; Record 2010). There is a rather large variety of scientific instruments exhibiting different properties and functions and some philosophers have tried to develop classifications in order to create systematicity and to better understand what different instruments do. In this section, I briefly review three current classifications of scientific instruments. The first two are classifications of scientific instruments in general and the third is concerned with instruments in biomedical engineering.

2.1 Current Classifications

Borrowing from Robert Ackerman (1985), Rom Harré (2003) refers to the total set of instruments available to a researcher as its "instrumentarium". This includes

[49] The classification of artifacts developed in this chapter is not meant to be exhaustive, i.e., it does not classify all cognitive artifacts in MBLs. Given the large variety of such artifacts in MBLs this would simply be unfeasible. A glance at any catalogue of scientific instruments shows that there are thousands of different instruments used in biology laboratories. The point I want to make is that the taxonomy I developed in chapter 4 is a useful tool for classifying cognitive artifacts in any given context.

instruments that are not present in the laboratory but can be ordered via the catalogues of instrument makers. Harré develops a set of categories within the instrumentarium "based on an analysis of laboratory equipment/world relationships" (Harré 2003, p. 25). His focus is thus not on how a researcher interacts with the instrument, but on how the instrument interacts and interfaces with the world. He starts by making a distinction between "instruments" and "apparatuses" and then makes further distinctions within those categories. Instruments "are causally related to some feature of the world in a reliable way", and apparatuses serve as a "working model of some part of the world" (Harré 2003, p. 26).

Two types of instruments are distinguished by Harré on the basis of the properties they measure. Instruments that measure primary properties do so in a format that somehow resembles the state of the measured property, whereas instruments that measure secondary properties do so in a format that does *not* resemble the state of the property. Thus, an electron-microscopic scan of a bacteria resembles its target and thus measures a primary property, whereas a thermometer measures the rapidity of molecular motion, not by resembling it, but by providing some other representational format. Harré continues by making a distinction between two kinds of apparatuses. First, material models that are "domesticated" versions of natural systems. These are artificially generated or cultivated versions of natural phenomena so that they can be studied in the laboratory, for example, a Drosophila (fruit flies) or E. coli (bacteria) colony[50]. Second, apparatuses that are integrated with the phenomenon under study, resulting in "apparatus-world complexes". These bring into existence what Harré refers to as Bohrian artifacts. Physicist Niels Bohr argued that certain instruments are not detached from the phenomenon of study but are sometimes part of it. This kind of apparatus brings into existence phenomena that do not exist in nature. An example of a Bohrian artifact that Harré puts forward is the isolation of sodium in metallic state by electrolysis. Metallic sodium does not occur in nature but only as part of an apparatus-world complex.

[50] Drosophila and E. coli are what biologists call model organisms. These are non-human species that are extensively studied to better understand certain phenomena. They are in vivo models and are used to examine human diseases when there may be ethical issues related to the research when performed with human subjects.

Davis Baird (2003, 2004) develops three categories of scientific instruments. Category membership in Baird's classification is based on the artifact's function. The first category consists of models whose function is to represent some target system. An example Baird mentions is Watson and Crick's physical model of DNA. The second category consists of instruments whose function is to "create a phenomenon" such as Michael Faraday's electromagnetic motor, Robert Boyle's air pump, and particle accelerators like the cyclotron. These instruments create experimental phenomena (i.e., rotary motion, vacuum, and fast moving particles) that are subsequently studied and measured. This brings us to the third category: measurement instruments. These include thermometers, spectrometers, calorimeters, voltmeters, and so on. Measurement instruments provide a researcher with information about some target system, for example, temperature, light, energy, electricity, and so forth.

The above two classifications are of scientific instruments in general. I now outline a classification of instruments used specifically in a biomedical engineering (BME) laboratory, which we have already seen in chapter 2. Based on a detailed ethnographic study of a BME laboratory, Nersessian et al. (2003, see also Kurz-Milcke, Nersessian & Newstetter 2004; Nersessian 2006) present a concise taxonomy of artifacts used in the laboratory. Their taxonomy has three categories: devices (e.g., a flow loop, bioreactor, or bi-axial strain), instruments (e.g., a confocal microscope, flow cytometer, or coulter counter), and equipment (e.g., a pipette, flask, refrigerator, or water bath). These three categories and the grouping of artifacts in those categories are developed by the BME researchers themselves. This was done during a research meeting with the members of the laboratory. On the basis of further ethnographic observations, Nersessian and her colleagues formulated working definitions of the categories. Devices are defined as engineered fascimiles that function as in-vitro models and sites of simulation. Instruments function as to generate measured output in visual, quantitative, or graphical format. And equipment assists with manual or mental labour.

2.2 Evaluation

The above classifications aim to classify what Ackerman (1985) has called the instrumentarium, i.e., *all* instruments and artifacts used by scientists, including those that do not exhibit cognitive functions. My aim here is more specific, as I am concerned only with *cognitive* artifacts in MBLs and will thus not discuss instruments that exhibit non-cognitive functions such as instruments that have as their function to create

experimental phenomena like, for example, Boyle's air pump, or instruments that have supportive roles such as water baths or refrigerators. In what follows, I first point out some similarities between these classifications, then some dissimilarities, and I end with some drawbacks. But note that these are drawbacks only relative to my specific aim.

In one way or another, all of the above classifications include a category of instruments that measure or detect some target system. Such instruments are of paramount importance for scientific research and are ubiquitous in all kinds of laboratories. In line with Peirce's view on representation, I will classify such instruments as indexical cognitive artifacts or indices. In chapter 3, we have seen that indices exhibit a direct causal connection to their target system. A thermometer, for instance, has a direct connection to the temperature: if the temperature changes, then the reading on the thermometer changes as well. Harré argued that what I call indices either represent their target by resembling it (i.e., by depicting primary properties) or by representing it in some other format (i.e., by depicting secondary properties). This links to the idea that a token representation can exhibit a combination of iconic, indexical, or symbolic properties. Thus a thermometer is an index that combines indexical and symbolic properties, whereas a confocal microscope is an index that combines indexical and iconic properties. All of the above classifications also include a category of models. Models typically (though not necessarily) display some kind of isomorphism to their target and are therefore members of the species of iconic cognitive artifacts or icons. Like indices, icons (including models) are essential to many sciences. In the next section, I elaborate in much more detail on different categories of cognitive artifacts in MBLs.

An important difference between these approaches is that Harré's classification is based on the instrument-world relationship, whereas the other classifications are based on the instrument's function. Otherwise put, Harré focuses on *how* an instrument does what it does, whereas Baird and Nersessian focus on *what* an instrument is used for. Further, Nersessian's taxonomy is distinct from the other classifications in that the categories are first devised by the researchers themselves and then further developed on the basis of ethnographic observations. So it is a much more bottom-up and practice-oriented method of classifying scientific instruments, as compared to the more conceptual methods of the other classifications.

The above classifications demonstrate that a large part of the instrumentarium consists of cognitive artifacts, as the majority of the examples concerned indices and icons, which are used to measure or model some target system. However, because current classifications do not focus on cognitive artifacts, but on instruments in general, they have overlooked some types of cognitive artifacts. That is, they have not included what I have referred to as symbolic cognitive artifacts and ecological cognitive artifacts. When theorizing about the genus of representational artifacts, I identified three species: icons, indices, and symbols. Symbols obtain their content and meaning, not through isomorphism or a direct causal connection, but through shared use, convention, and logical rules. Examples in scientific practice include written language, numbers, equations, mathematical formalisms, formulae, and tables. Like icons and indices, symbols are essential to scientific practice. Examples of artifacts containing symbols in scientific practice include manuals, protocols, textbooks, articles, notebooks, patents, calculators, or computing systems. Such artifacts have quite important roles in scientific research and a classification of cognitive artifacts in MBLs should include such artifacts.

Similarly, what I have called ecological cognitive artifacts are important for scientific practice, too. These are artifacts that do not provide representational information (i.e., information about the world), but ecological information (i.e., the world as information). I distinguished between two species of ecological artifacts, namely, those that exhibit spatial and dynamic properties. Examples of spatial ecological artifacts include consistently leaving car keys on a certain spot in your apartment so that you know where they are, putting an article you have to read on top of the pile on your desk, leaving a book open and turned upside down so that you know where you have stopped reading, or tying a string around your finger as a reminder. These artifacts contribute to performing some cognitive task by exhibiting ecological information, i.e., they do not have targets. Rather, the idea is that by putting artifacts in specific locations that are either deliberately usual or deliberately unusual, we encode task-relevant information into the artifact and its location, such that when we encounter the artifact in that location it typically prompts a memory. Examples of dynamic ecological artifacts include the artificial rotation of zoids in playing Tetris and the rearranging of letter tiles in playing Scrabble. New information emerges when manipulating these structures, implying that their cognitive function supervenes on a manipulable structure. Given the large amount of instruments researchers (in MBLs) use and the complexity of their

cognitive tasks, they are likely to structure their work-environment such that it facilitates their ongoing cognitive task. This "intelligent use of space" (Kirsh 1995) is most likely quite common in laboratories, including MBLs, but has not yet been addressed in the literature and has not been included in classifications of instruments.

3. Classifying Cognitive Artifacts in Molecular Biology

In chapter 3, I argued for an information-centered approach to taxonomizing cognitive artifacts. This approach entailed focussing on the particular informational properties of cognitive artifacts and categorizing them on the basis of those properties, rather than on the basis of the properties or goals of the agents that design, select, improvise, or use them. I developed this approach because cognitive functions supervene on informational properties, i.e., cognitive artifacts are used in virtue of the information they contain. In the taxonomy, I first distinguished between representational and ecological artifacts, and then made further distinctions within those genera. In the case of representational artifacts, I distinguished between artifacts exhibiting iconic, indexical and symbolic properties. And in the case of ecological artifacts, I distinguished between those that exhibit spatial and dynamic properties. Thus at the species level of the taxonomy there are five categories which should not be seen as mutually exclusive because they sometimes overlap, but typically one informational property is predominant. In this section, I classify cognitive artifacts used in MBLs in terms of this taxonomy. However, as stated, I do so based on their predominant informational property and so artifacts in one category may display a combination of informational properties.

3.1 Representational Artifacts

3.1.1 Icons

Icons are defined by exhibiting relevant isomorphism with their targets. Before discussing some examples of icons in MBLs, it is helpful to point out that there are at least two different kinds of isomorphism. First, there may be structural isomorphism between icon and target. Watson & Crick's model of DNA is an example of a structurally isomorphic icon. Second, there may be sequential isomorphism in which the icon represents the order of steps in a process or mechanism. A diagram depicting the order of the steps in a biochemical process (e.g., DNA → mRNA → protein) is sequentially isomorphic in that it represents not some physical state but the sequence of steps in a

process. Note that this particular example is also symbolic as the acronyms DNA and mRNA and the word protein obtain their meaning through shared use, rules, and convention.

There are a variety of artifacts founds in MBLs that exhibit isomorphism to some target system. The most prominent of these are photographs, diagrams, and scientific models. Photographs are made by light falling on a light-sensitive surface, typically resulting in structurally isomorphic images. Molecular biology extensively makes use of photographic techniques often combined with microscopic techniques. Sometimes photos are taken of organs or other similar sized biological structures in order to document the size, colour, and structure of the object in question. But more often photos are taken of microscopic structures such as clusters of cells, individual cells, cell organelles, DNA molecules, proteins, cell receptors, vesicles, and so on. This is typically done with microscopes, of which there are various types, including optical microscopes, electron microscopes, scanning probe microscopes, digital holographic microscopes, and other types of microscopes.

Microscopes as instruments exhibit a combination of indexical and iconic properties. There is a direct causal connection between what is seen through a microscope and the specimen, which is usually mediated by an electronic screen. But after a photo is taken with it, the indexical part of the representational system is gone and the photo becomes purely iconic. Some molecular biologists are interested in cell processes or mechanisms that take place over time such as, for example, cell division. In such cases, either a series of photographs or videos are taken. In either case, a series of photos or a video are both structurally and sequentially isomorphic to their target, in that they resemble the physical structure of dividing cells and represent the sequence of steps in the process.

Of all the artifacts used in science, diagrams and models have received the most attention from philosophers (Hughes 1997; Magnani, Nersessian & Thagard 1999; Giere 2002a; Klein 2003; Knuuttila 2005, 2011; Charbonneau 2013; Friggs & Hartmann 2012; Toon 2012). A helpful way of thinking about the difference between models and diagrams is that models are defined by their function, whereas diagrams are defined by their format. Both models and diagrams function as to represent some target, but models can do so in a variety of formats, including two-dimensional (e.g. diagrams), three-dimensional (e.g. physical models), and four-dimensional (e.g. computer

simulations). Diagrams, by contrast, are defined by representing their target in a certain format, i.e., in a two-dimensional graphical representation of information that typically displays some kind of isomorphism to its target. Examples include tree diagrams, network diagrams, flow charts, Venn diagrams, histograms, pie charts, population density maps, and exploded views of objects or structures. Models can thus be seen as a larger category and diagrams as instances of that category[51].

There are countless models used in molecular biology that are isomorphic to their targets (Sheredos, Burnston, Abrahamsen & Bechtel in press; Perini 2005, 2013)[52]. Examples of structurally isomorphic models in molecular biology include two-dimensional and three-dimensional models of the structure of biomolecules, such as proteins, DNA, lipids, and so forth. A key example is Watson and Crick's physical model of the double helix of DNA. Examples of sequentially isomorphic models are representations of processes and mechanisms such as, for example, cell division, gene expression, metabolism, endocytosis, and so on. Usually such icons contain arrows to indicate steps in a process. Key examples here are a diagram of the Krebs circle, containing many arrows to indicate the numerous steps in the metabolic process, or a diagram of circadian oscillators in mammals (Bechtel & Abrahamsen 2013). Another key example of sequentially isomorphic models are computer simulations of biochemical processes and mechanisms such as, for example, protein folding. Given the benefits of computer simulations like information storage capacity and processing speed and capacities, they are becoming increasingly more popular as modelling tools (Charbonneau 2013).

---

[51] However, note that some models may display isomorphism to their target, while others do not. For example, Watson and Crick's physical model of the structure of DNA is isomorphic to its target system, whereas a mathematical model of quantum gravity is not isomorphic to its target system. Therefore, some models are predominantly iconic and others are predominantly symbolic. In this chapter, I only discuss models that are predominantly iconic.

[52] It has been claimed that biologists use more diagrams than do other scientists. Laura Perini (2005) has argued that a possible explanation why diagrams are so often used in (molecular) biology is because biology is concerned with functional explanations rather than mere causal explanations. And diagrams are particularly well-suited to represent functions of components of systems.

3.1.2 Indices

Indices are defined by exhibiting a direct causal connection between the representational state of the index and its target system. Molecular biologists use a large variety of indices, including graduated cylinders to measure the amount of liquid, thermometers to measure temperature, pipettes to add a precise amount of liquid, hemocytometers to count to number of cells in a solution, pH meters to measure the acidity of a solution, scales to measure the weight of some substance, pycnometers to measure fluid density, spectrometers to measure the amount of light passing through a solution or substance, etc. As was also recognized by Harré, Baird, and Nersessian, indices are essential to many scientific practices, including those in molecular biology. Given that science investigates the world empirically, the best way to do so is by probing, measuring, and mapping it with the aid of indices. Moreover, given that molecular biology deals mainly with microscopic structures, which are inherently invisible to the human eye, there is only one way to empirically obtain information about such target systems and that is with indices.

An important research method for molecular biologists is gel electrophoresis, which is an experimental method for the separation and analysis of large biomolecules, such as proteins or DNA, on the basis of their size. In this method, biomolecules migrate through a gel with a certain porosity. One side of the gel is charged positively and the opposite side is charged negatively. Given that biomolecules are charged, they will be pulled towards either side of the gel. Large biomolecules encounter more resistance and will thus migrate shorter distances than smaller molecules. Thus in this method, the location of the molecules in the gel is an index for their size.

On the basis of the above example, one can distinguish between dynamic and static indices. The representational state of dynamic indices automatically changes when its target system changes. For example, when the temperature changes, the reading on a thermometer changes as well. This also happens with pH meters and the acidity of a solution, spectrometers and the amount of light passing through a solution or substance, and other indices. By contrast, gel electrophoresis can be seen as a static index, as its representational state does not change when its target changes. After the electrophoresis has been performed, a static representational structure is created. Due

to the properties of this method, one cannot change the target system in the same way as is the case with dynamic indices.

3.1.3 Symbols

Symbols obtain their meaning and content through shared use, agreement, and logical rules. Typical examples are natural and artificial languages, numbers, mathematical systems, tables, equations, and symbols in scientific formulae. Symbols and symbolic structures are ubiquitous in MBLs, as they are found in manuals for equipment, textbooks, scientific articles, and laboratory notebooks. They are also found on indices such as pipettes, thermometers, voltmeters, and other indices that indicate the quantity of their target systems in numbers. A paradigm symbol in molecular biology is the period table of elements, which is a table of the chemical elements organised on the basis of their atomic numbers. There is no isomorphism or direct causal connection between the table and its target, i.e., the order in which the elements are presented is purely based on logical rules and agreement.

Another prominent symbolic cognitive artifact is the laboratory notebook. Researchers in MBLs have a personalised laboratory notebook in which they document their hypotheses, their experimental procedures and outcomes, and observations made during the experiment. The author of a biology textbook writes:

"A laboratory notebook is one of a scientist's most valuable tools. It contains the permanent written record of the researcher's mental and physical activities for experiment and observation, to the ultimate understanding of physical phenomena. The act of writing in the notebook causes the scientist to stop and think about what is being done in the laboratory. It is in this way an essential part of doing science" (Kanare 1985, p. 1).

The majority of information in laboratory notebooks is symbolic in nature (e.g. language, equations, tables, calculations). For this reason, I classify the notebook as a predominantly symbolic cognitive artifact, but they may also contain iconic representations such as diagrams or sketches.

3.2 Ecological Cognitive Artifacts

3.2.1 Spatial Ecological Artifacts

Spatial ecological artifacts obtain their cognitive function in virtue of physical-spatial structures. More specifically, by intentionally putting artifacts in usual or unusual places we encode task-relevant information into the artifact and its location. I may, for instance, always leave my car keys on my hall table, put a rented DVD on my desk as a reminder to bring it back to the video store, or structure my work-environment (e.g. a kitchen or workshop) such that the location of the artifacts facilitate the cognitive task I am performing. Such "intelligent use of space" (Kirsh 1995) also happens in MBLs. A researcher may, for instance, intentionally leave an alcohol bottle on the laboratory bench as a reminder to clean and disinfect the bench after doing an experiment. Researchers may also structure their work-environment in more complex ways. When performing an experiment, they use a variety of instruments like, for example, pipets, test tubes, beakers, flasks, racks, stopwatches, heaters, etc. Experimental procedures require that these instruments are used in a certain order. Novice researchers often use a laboratory notebook in which the experimental procedure is outlined, which helps them to remember the correct order of the steps in the experiment. But more experienced researchers may have developed strategies that make the notebook (partly) superfluous.

Consider the following example. When performing a DNA isolation procedure, these are the main steps: (1) break cells open by adding lysozyme, (2) remove membrane lipids by adding a detergent, (3) remove proteins by adding protease, (4) remove RNA by adding RNase, and (5) purify DNA by ethanol precipitation. Given the order of these steps, it makes sense to organize the workplace such that the lysozyme is located close to the experimenter, as this is the chemical she needs first. Likewise, the detergent should also be located quite close as this is the second chemical she needs in the overall procedure, and so on. How the workbench and the artifacts are organized is important for performing the cognitive task at hand. For this reason, researchers prepare the laboratory bench before they begin the experiment. They will most likely also create locations on the workbench that are reserved for artifacts that have been used and are no longer needed. Thus when the first step is performed, the bottle containing lysozyme will be put at a spot that indicates it has been used. For example, the chemicals needed for the experiment may be put on her right hand side in the same

order in which they are used in the experiment, and the chemicals she has used may be put on her left hand side, thereby creating spatial boundaries that help the researcher in performing the experiment.

3.2.2 Dynamic Ecological Artifacts

Dynamic ecological artifacts obtain their cognitive function in virtue of a manipulable physical structure. Rearranging letter tiles in playing Scrabble, for example, creates new combinations of letters which ideally prompts word recall. Likewise, artificially rotating zoids when playing Tetris outsources mental rotation and helps to perform the task. Rearranging letter tiles or zoids is a dynamic and ongoing process in which physical structures are manipulated as to create new task-relevant information. I was unable to find any examples of dynamic ecological artifacts in MBLs in the literature. I do, however, think it is likely that such artifacts occur in MBLs, but this hypothesis can only be tested by a more elaborate literature review or by empirical observations (see also next chapter).

4. Distributed Cognition in Scientific Practice

Before I conceptualize how and how deeply some of the above classified cognitive artifacts and their users are integrated into a larger cognitive system, I survey a small body of literature concerning distributed cognition and scientific practice. I specifically look at the work of Ronald Giere, P.D. Magnus, and Nancy Nersessian.

4.1 Ronald Giere

In a series of articles, Giere has put forward distributed cognition theory, as developed by Hutchins (1995), to better understand certain aspects of scientific research. For example, reinterpreting Bruno Latour's (1986) paper, *Visualisation and Cognition: Thinking with Eyes and Hands*, Giere and Moffatt (2003) say that a chemist and the symbols in a chemical formula he is manipulating are constituting a distributed cognitive system. They write:

"Understood in terms of distributed cognition, these formulas are external representations that form part of a distributed cognitive system for exploring possible reactions in organic chemistry. The cognitive process of balancing an equation does not take place solely in the head of some person, but consists of interactions between a person and physical, external representations" (Giere and Moffatt 2003, p. 4).

Giere and Moffatt point out that distributed cognitive systems in scientific practice, like the one described above, are important as they make possible the acquisition of knowledge that no single agent or group of agents could do without cognitive artifacts. Elsewhere Giere (2002a) investigates how scientific models can be part of distributed cognitive systems. He says that models are not just aids for cognition, but are, when properly manipulated, part of a distributed cognitive process. The point Giere makes is that human agents cannot store, let alone manipulate, complex scientific models in their head. They need to create and manipulate external representations that complement the shortcomings of the pattern matching abilities of embodied brains. When talking about physical models, such as Watson and Crick's model of the double helix of DNA, Giere says:

"Physical models provide what is probably the best case for understanding model-based reasoning as an example of distributed cognition. Here it is very clear that one need not be performing logical operations on an internal representation. It is sufficient to perform and observe appropriate physical operations" (Giere 2002a, p. 9).

He further points out that "by fiddling with the physical model so as to fit the base pairs inside a double-helical backbone, they came up with the right structure" (Giere 2006a, p. 104). So here we have two cooperating embodied agents that manipulate a physical model, thereby constituting a distributed cognitive system. Giere makes an interesting claim about the affordances of models exhibiting different kinds of representational formats. He claims that Watson and Crick with their physical three-dimensional model were able to find the correct structure of DNA, whereas Rosalind Franklin who only had some two-dimensional X-ray photographs and hand-drawn diagrams was unable to find the structure of DNA. There may have been other factors involved in the discovery of the structure of DNA, but Giere nonetheless is onto something here. The specific informational properties and affordances of the model are important for performing the cognitive task. So scale models in this case are better for performing this cognitive task as compared to sketches on paper, because they complement onboard cognitive systems in a more effective way. Perhaps this is so because the human organism has evolved to deal with a three-dimensional environment and is therefore better equipped to understand three dimensional structures. Watson and Crick's physical model affords different cognitive functions as compared to a sketch of the same structure. Their

physical model and a sketch on paper may be informationally equivalent in that they contain the same informational content, but they are computationally non-equivalent because they afford different interactions, resulting in different situated cognitive systems.

Giere also theorizes about larger distributed cognitive systems. For example, when talking about the Indiana University Cyclotron Facility, he says that

"It is particularly enlightening to think of the whole facility as one big cognitive system comprised, in part, of lots of smaller cognitive systems. To understand the workings of the big cognitive system one has to consider the human-machine interactions as well as the human-human interactions" (Giere 2002b, p. 8).

He then goes on to list some of the components of this big cognitive system, including the accelerator, detectors, computers, and the people actively working on the experiment. The social structures in the research facility are also an important organisational aspect of the overall cognitive system. The researchers monitoring the data acquisition typically are staff members with PhDs, those tending to the detectors may or may not have PhDs, and those that keep the accelerator in tune are technicians without PhDs. So one's academic status and level of education determines one's role in the overall system. Likewise, Giere writes that the Hubble Telescope and the researchers using it can be seen as one distributed cognitive system. A single experiment may include fifty to five hundred people and countless computing systems processing the data integrated into one cognitive system. In a similar fashion, Giere (2002c) reinterprets the work of sociologist of science Karin Knorr Cetina (1999). Knorr Cetina performed long-term observations at two different laboratories: (1) the European Centre for Nuclear Research (CERN) and (2) a MBL. Giere argues that CERN can be seen as a distributed cognitive system. He writes that when an experiment is performed using the Large Hadron Collider and the ATLAS detector, which may involve up to a hundred researchers, "we attribute the cognitive capacity to acquire the desired data to the whole system, people plus machines organized in an appropriate way. The cognition is in this way distributed" (Giere 2002c, p. 5).

4.2 P.D. Magnus

Like Giere, P.D. Magnus (2007) applies Hutchins' distributed cognition framework to explore scientific practice. He argues that the distributed cognition framework only applies to activities that perform a task with a clear goal or function such as, for example, navigating a ship or landing an airplane. Because it is difficult to attribute a task and goal to science in its entirety, i.e. on a global level, it is not possible to use the distributed cognition framework when explaining science on this high level of abstraction. Therefore, as Magnus points out, views on science that try to characterize it on a global level such as Robert Merton's sociology of science and Philip Kitcher's ideas on well-ordered science are incompatible with a distributed cognition view on science. What we need to do, instead, is identifying tasks and goals on a local level, i.e., at the level of specific scientific projects. Magnus puts forward an example of Nersessian (2005, see also next section) who has studied the use of cognitive artifacts in a biomedical engineering laboratory. The task of that laboratory can be seen as understanding and simulating blood vessels. Given that the task can be specified, the work in the laboratory is an instance of a distributed cognitive system.

4.3 Nancy Nersessian

We have already briefly encountered the work of Nersessian in chapter 2 and in this chapter. Here I review her research on distributed cognition in a biomedical engineering (BME) laboratory. An important difference between her work and that of Giere and Magnus is that they theorize on distributed cognition in science on a conceptual level, whereas Nersessian's work is based on extensive, long-term empirical research. Nersessian points out that BME labs differ from the paradigm cases of distributed cognitive systems that Hutchins (1995a, 1995b) has analysed. She argues that pilots in a cockpit or navigators on board of a ship face problems that change over time, but the nature of the artifacts and the knowledge that the agents bring to the situation remain largely stable. Such systems are thus dynamic but *synchronic*, i.e., the functional properties of the cognitive artifacts remain largely the same over time. By contrast, a BME lab can be seen as an evolving distributed cognitive system, i.e., the knowledge of the scientists and the functional and informational properties of the instruments and artifacts change over time. So the BME lab as a distributed cognitive system is dynamic and *diachronic*. Both researchers and artifacts change in relation to each other, so they have relational trajectories of change. Due to the diachronic nature of the laboratory,

Nersessian develops an approach she refers to as a cognitive-historical analysis (Nersessian 2005, 2008, 2009). She writes:

"Cognitive-historical analysis enables following trajectories of the human and technological components of a cognitive system on multiple levels, including their physical shaping and reshaping in response to problems, their changing contributions to the models that are developed in the lab and the wider community, and the nature of the concepts that are at play in the research activity at any particular time. As with other cognitive-historical analyses, we use the customary range of historical records to recover how the representational, methodological, and reasoning practices have been developed and used by the BME researchers" (Nersessian 2005, p. 17).

An important element in this approach is to first identify the relevant cognitive artifacts and then study how they are used and how their properties change over time. This is done by a combination of ethnographic observations and interviews.

The relation between researcher and cognitive artifact is characterized as a *cognitive partnership*. "These partnerships provide the means for generating, manipulating, and propagating representations within the distributed cognitive systems of this research laboratory" (Nersessian 2005, p. 24). Moreover, this partnership transforms both researcher and artifact. A newcomer in the lab, for example a new PhD student, may see many of the cognitive artifacts as mere objects, but over time when s/he interacts with the artifacts and learns their affordances and functions, the capabilities of the student are transformed. Likewise, during their use, some cognitive artifacts are adjusted or re-engineered. "Re-engineering is possible because the researcher with a developed partnership appropriates and participates in the biography of the device" (Nersessian 2005, p. 24). Elsewhere she writes: "We use the notion of cognitive partnership to characterize the special relations that develop among researchers and between them and simulation devices (as opposed to other important artifacts such as the pipette or confocal microscope) in the course of learning in the lab" (Nersessian 2009, p. 741). Although pipettes and confocal microscopes generate or contain external representations and can thus be considered as cognitive artifacts according to Nersessian's definition, they do not enter into cognitive partnerships with the agents that use them. Cognitive partnership thus concerns only diachronic situated cognitive systems. Finally, Nersessian writes that cognitive artifacts generate, manipulate, or

propagate *representations*. So on her view, only artifacts exhibiting representational properties are part of distributed cognitive systems (see also chapter 2, section 2.2.5).

4.4 Evaluation

The above approaches are perceptive and I am largely in agreement with them. There are, however, a number of things that I believe need further evaluation. Giere develops a view on distributed cognition that is in line with the motivation for complementarity-based extended mind theory. Consider the following quote:

"Most models in science, even in classical mechanics, are too complex to be fully realised as mental models. Not even authors of science textbooks can have in their head all the details of the models presented in their text. Rather, the details of these models are reconstructed as external representations when needed. These reconstructions typically take the form of equations or diagrams. What scientists have inside their skins are representations of a few general principles together with bits and pieces of prototypical models. They also possess the skills necessary to use these internal representations to construct the required external representations that are then part of an extended cognitive system" (Giere 2002b, p. 10).

This quote is in line with the complementary-based extended mind theory that has been developed in the previous chapter. Because the human brain cannot store large complicated scientific models, it only stores the most important aspects of the model and also knows how to use that information to reconstruct the entire model externally. The externalized information and the embodied skills to create that information thus complement the shortcomings of the brain. I agree with Giere's motivation, but he does not provide any information about why scientists and their diagrams or models constitute distributed cognitive systems, i.e., he does not look at the criteria or conditions for cognitive distribution. He argued that, when properly manipulated, models are part of distributed cognitive systems, but does not explain what properly manipulated means.

This lack of criteria may be one of the reasons why some of Giere's examples are perhaps somewhat unsupported. For instance, he argued that entire research facilities such as CERN and the Indiana University Cyclotron Facility, including all the people performing experiments, constitute one large distributed cognitive system. It is one

thing to claim that entire research facilities are one distributed cognitive system, but it is quite another thing to explain in detail how thousands of researchers interacting with instruments and with each other to constitute such a system. By contrast, Hutchins' study of ship navigation consisted of a detailed ethnographic study, investigating the properties and functions of and information flow between each component of the distributed cognitive systems on the Palau jointly doing the navigating. But Giere does not support his grand claims with any detailed descriptions or empirical findings, nor does he provide any criteria for determining when a cognitive system is distributed, resulting in an account that is perhaps too liberal. A final shortcoming of Giere's approach is that he exclusively focusses on representational artifacts and neglects what I have referred to as ecological artifacts[53].

I am sympathetic to Magnus' main claim. His article is an exercise in determining whether a distributed cognition approach is feasible for explaining science. Basically, his point is that if we want to use a distributed cognition approach to studying science, we should look at specific situated cognitive systems and not at science in its entirety. I agree with this rather general point, but after having assessed the usefulness of distributed cognition for better understanding science, a number of concrete examples would have been helpful. He writes: "There is a wealth of literature on scientific instrumentation and social networks in science, much of which could be framed in the d-cog idiom" (Magnus 2007, p. 7). This is true, but unfortunately no examples are given by Magnus in which he actually uses a distributed cognition idiom to theorize about scientific instruments and social networks. His approach thus remains highly speculative. Lastly, Magnus works with a parity-like criterion for cognitive distribution:

"We can say that, in order for an activity to count as d-cog, the task must be such that it would be cognitive if the process were contained entirely within the epidermis of one individual. That is: It is the sort of task that could be carried out in a mind. So an activity is d-cog if (1) the task is such that it would count as cognition if it were carried out entirely in a single mind or brain, and (2) the process by which the task is carried out is not enclosed within the boundary of a single organism" (Magnus 2007, p. 4-5).

[53] See also Vaesen (2011a) for a critique on Giere's view on distributed cognition.

Magnus thus claims that cognition can only be distributed when the external component functions in the same way as an internal state or process. In the previous chapter, I argued that parity considerations are unproductive when trying to argue for extended minds or distributed cognitive systems (see also Cheon 2013 for a critique of Magnus). Rather, we should understand the properties and functions of external components in situated cognitive systems as complementary to internal states and processes, not as isomorphic. By using parity-like considerations, Magnus misses the point that external components do not exhibit identical properties and functionalities as internal states and processes.

Of the three approaches outlined above, I find Nersessian's the most promising. However, in her analysis she, like Giere, focuses exclusively on artifacts containing external representations (see also chapter 2, section 2.2.5). As I have argued in previous chapters, artifacts can also contribute to performing cognitive task by exhibiting non-representational or ecological information. Thus, neither Nersessian, Giere nor Magnus address what I have called ecological cognitive artifacts. In order to provide a more complete picture of the variety of situated cognitive systems in MBLs (or any other laboratory or setting), we need to include ecological cognitive artifacts in the analysis. Furthermore, I think Nersessian's empirical method can be enriched by incorporating elements from the multidimensional framework developed in the previous chapter. In order to better understand the "cognitive partnership" between researcher and artifact, it is helpful to utilize my multidimensional framework.

For example, Nersessian's emphasis on the diachronic aspect of distributed cognitive systems can be further conceptualized by the dimensions of trust, procedural transparency, informational transparency, individualization, and transformation. These dimensions are meant to capture the diachronic nature of situated cognitive systems. Over time, researchers learn to trust the informational properties of the artifact. When they get more skilful in using and manipulating the artifact, the procedural transparency increases and they find it easier to interpret the information of the artifact and so the informational transparency increases as well. During an experiment, the artifact is individualized in terms of the epistemic goals of the experiment. The informational properties of the artifact are transformed and, depending on the amount of use, the artifact may also have transformed the onboard representational and cognitive system of its user, i.e., the researcher learns to think in terms of the

informational properties of the artifact. In the next section, we will look in more detail at the integration between researchers and cognitive artifacts in terms of the dimensions developed in the previous chapter.

4.5 Top-Down and Bottom-Up Approaches to Distributed Cognitive Systems

Knorr Cetina's (1999) research at a MBL showed that there are two organisational levels. The higher level concerns the whole laboratory which is managed by the head of the laboratory, whereas the lower level concerns individual researchers working on their own projects. MBLs are thus different from physics laboratories in which there is typically only a one-level structure, i.e., one in which there is a large-scale cooperation between researchers, although this may not be true for all physics and MBLs. Giere argues that this has effects for the kinds of distributed cognitive systems that exist within the lab, which do not contain many researchers and a variety of artifacts and instruments, but are composed of individual researchers and their cognitive artifacts. I think this is an important insight. Different scientific laboratories exhibit different kinds of situated or distributed cognitive systems, depending on the kind of research that is performed in the lab and the social cooperation and cognitive artifacts that that research requires. Physics, according to Cetina Knorr, is more cooperative in nature than molecular biology, which has consequences for situated cognition. I want to add to this observation that I think it may not be generalizable to all physics and MBLs. There are clearly MBLs that do exhibit larger distributed cognitive systems in the sense advocated by Giere. For example, the Human Genome Project (HGP) was a large-scale, long-term, multi-university research project comprising of thousands of researchers and countless instruments. There are, likewise, many projects in molecular biology that are large-scale and involve many researchers and even research institutes (Parker, Vermeulen & Penders 2010).

On the basis of Knorr Cetina's work, one may distinguish at least two approaches to studying distributed cognitive systems in scientific laboratories. One can start by identifying a large-scale system such as CERN or the HGP and then try to identify all the relevant components and see how information trajectories are established and how all the components are integrated into one large distributed system. This may be called a top-down approach. Given that there are sometimes more than a hundred researchers involved and numerous artifacts and instruments, the information trajectories and the overall system are so complicated that no single ethnographer can observe such a

system. If one chooses a top-down approach it is only feasible if the system is sufficiently small to observe or that there are sufficient ethnographers to study it. If a distributed cognitive system is comprised of more than a hundred researchers interacting with highly complicated machinery, it is in practice quite hard, if not impossible, to observe how the overall system and its components operate. This is presumably why Giere, in some of his examples, theorizes at a very general level of explanation without giving specific details of the composition and integration of the large-scale systems he is talking about.

One may also start at identifying small-scale systems, comprising of an individual researcher and the artifact(s) he or she is interacting with. If that small-scale system is part of a larger one, it may be possible to work your way up and include more components in the analysis. The approach one chooses clearly depends on the kind of situated system one is interested in. Sometimes only a bottom-up approach is possible as there may not be a larger distributed system, which may be the case in small laboratories. But, when the system in which one is interested is sufficiently small, one can adapt a top-down approach. And sometimes a combination of the two approaches is the best option. In the next section, a bottom-up approach is used to conceptualize how researchers interact with their instruments and how they are integrated into a larger cognitive system.

5. Conceptualizing Integration: Case Studies

Having classified some of the cognitive artifacts in MBLs and reviewed current literature on distributed cognition and scientific practice, let us now see how and how deeply artifacts and the cognitive systems of their users are integrated into distributed cognitive systems. I now present four concise case studies: (1) computer simulations of protein folding, (2) pH-meters, (3) laboratory notebooks, and (4) organised workplaces. These are instances of cognitive artifacts at the species level of the taxonomy. Computer simulations are predominantly iconic, pH-meters are predominantly indexical, lab notebooks are predominantly symbolic, and organised workplaces are structural ecological artifacts. Because I was unable to find an example of a dynamic ecological artifact in MBLs, there is only one case study on ecological artifacts. I have chosen these particular cognitive artifacts because they are ubiquitous in molecular biology practice and to the best of my knowledge have not received attention from a

cognitive perspective. Before doing these case studies, it is helpful to briefly recap the dimensions: epistemic action and information flow, speed of information flow, reliability, durability, trust, procedural transparency, informational transparency, individualization, and transformation. Where relevant, each dimension is quantified in three values: low, medium, and high. Note that determining these values is not an exact science and involves judgment on my part. It is, for example, quite difficult to determine how much an artifact has transformed one's representational capacities.

5.1 Representational Artifacts

5.1.1 Computer Simulations

In chapter 4, we have seen that Peirce argued that any representational system is inherently and irreducibly triadic in that it always involves an agent, a representation, and a target. For analytical purposes, I decomposed this triadic relation into two dyadic relations; one between the agent and representation, and one between the representation and target. A similar approach to representational systems occurs in literature on models in philosophy of science. Some theorists focus on how models represent their target. This approach is referred to as the semantic view and an important topic in that approach is *how* models can represent their target (Friggs & Hartmann 2012). Other theorists focus on the relation between the user and model and stress that models are physical artifacts with certain affordances and functions (Morrison & Morgan 1999a, 1999b; Knuuttila & Voutilainen 2003; Knuuttila 2005, 2011; Knuuttila & Boon 2011; Vorms 2011, 2012). This approach is sometimes referred to as the pragmatic view. Tarja Knuuttila (2011) points out that "From this perspective models function as external tools for thinking, the *construction* and *manipulation* of which are crucial to their epistemic functioning" (Knuuttila 2001, p. 263, original italics). Thus, on this view, models are seen as cognitive artifacts whose affordances and functions are used in performing cognitive tasks related to scientific reasoning such as, for example, inference-making.

Nersessian (1999, 2002) and others (Magnani, Nersessian & Thagard 1999; Magnani & Nersessian 2002) have referred to this as model-based reasoning. An important point regarding model-based reasoning is that we learn about the model and its target not by merely observing the model but by actively manipulating it. Although this point has been noted by several theorists (Baird 2003; Knuuttila 2005; Vorms 2012; Charbonneau 2013), none of them provide an overall theoretical account of the interactive and

integrative process. Situated cognition theory in general and the multidimensional framework developed in this thesis in particular, provide an excellent theoretical framework for explaining the cognitive interactions between user and model.

When models are static two-dimensional or three-dimensional structures, researchers have to mentally visualize the changes the target system undergoes. Benjamin Sheredos et al (in press) argue that "Such mental simulation lacks quantitative precision and can be highly fallible. A researcher may overestimate the capabilities of a component part or neglect important consequences of a particular organisation, such as how it might alter another part" (Sheredos, Burnston, Abrahamsen & Bechtel in press, p. 6). It is precisely because our internal cognitive systems are limited that researchers use models to complement these limitations (see also Giere 2002b). Scientists use models for the same reason that artist's sketch, i.e., to complement the limitations of mental imagery (Van Leeuwen, Verstijnen & Hekkert 1999; Clark 2001). To better understand how particular models complement and are integrated into the cognitive systems of their users, I now take a closer look at computer simulations of protein folding.

After a protein is synthesized by a ribosome in a cell, it starts as a long chain of amino acids referred to as its primary structure. This primary structure first coils into a number of secondary structures such as alpha-helices and beta-sheets, which in turn fold into the eventual three-dimensional tertiary structure (Teter & Hartl 2005). A protein's tertiary structure determines its function and is therefore important to better understand. The problem molecular biologists face is how a long chain of amino acids eventually results in the tertiary structure of the protein. This folding process is quite complicated, not yet fully understood, and very difficult to study in vivo (within cells). For these and other reasons (see also Peck 2004), molecular biologists build computer simulations of the folding process. Such simulations are both structurally and sequentially isomorphic to their target systems in that the simulation of the folding process is structurally and sequentially similar to the actual folding process. They are thus four-dimensional representations, but note that their structure is much simplified and the folding process is slowed down significantly. Two reasons for focussing on computer simulations is because they are becoming increasingly popular in (molecular) biology (Winsberg 2013), and have received very little attention from a cognitive perspective. Paul Humphrey (2004, 2009) argued that computer simulations are philosophically interesting, both for epistemological and methodological reasons, but

did not address the cognitive aspect of the relation between simulation and researcher, which I will do below.

There are two main stages in the simulation process: building the simulation and running it. In most cases, the amino acid sequence of a protein is known and used as input for the simulation. Amino acids have certain properties that determine how they will interact with other amino acids. The software program contains all the necessary information to simulate the interactions between different amino acids such as information about bonds, angles, torsion angles, Lennard-Jones interactions, and electrostatic interactions. The epistemic actions in the building stage of the simulation mainly concern data input into the software program. When the simulation has been successfully built, it is run and visualized on a computer screen. The in vivo folding process takes milliseconds, but the simulation can be run as slow as a modeller wants. Typically the simulation is slowed down a hundred to a thousand times, in order to scrutinize each step in the folding process. Also, different parts of the protein are indicated with different colours. Alpha-helices may be indicated as blue, beta-sheets as yellow, the C-terminus as red, and so on. When modellers perceive and interpret the simulation, they use gestures and bodily actions as cognitive artifacts, i.e., as human-made information-bearing structures used for performing cognitive tasks. For example, when interpreting experimental data and simulations, researchers point at salient elements in the simulation and make gestures towards it to aid the interpretation process (Alec and Hutchins 2004; Becvar, Hollan & Hutchins 2005, 2007).

In line with Giere (2002b), I think that simulations are clear cases of cognitive artifacts that complement the shortcomings of embodied brains. Trying to mentally visualize a static three-dimensional structure such as a protein is already difficult, if possible at all, but trying to mentally visualize the folding of a protein over time, is for most people impossible. Computer simulations of the folding process thus complement the shortcomings of the information-storage and information-processing capacities of embodied brains. During a simulation, the angle of the overall protein and speed of the simulation can be altered, as to optimize perceptual access to salient properties of the folding process. One can also zoom-in on a particular part of the protein. Thus, in terms of information flow, a modeller first offloads information onto the computer system, then the computer creates the simulation with which a modeller interacts in an ongoing and iterative manner, thereby creating a reciprocal information-flow structure.

The time it takes to program the simulation is substantial and can take a few hours up to several days, depending on the complexity of the protein and how much is known about its tertiary structure. As a result, information offloading speed is low to medium. However, once the simulation has been programmed, the informational uptake is fast, as one only has to perceive the folding of the protein, which typically is slowed down to a few seconds up to a minute. It is fairly easy to determine how much information is offloaded onto the program, but it is much more difficult to determine how much information is taken onboard. A single simulation contains a substantial amount of information, as most amino acids interact with other amino acids, forming secondary structures, which then interact to form the tertiary structure. So there are many interactions and steps in the overall folding process, which are displayed in a few seconds up to minutes. It may not always be possible to determine what and how much information is actually perceived and processed by a modeller, but it is safe to say that the intake speed is higher than the offloading speed.

Access to the simulation is reliable as it is run on computer systems that are typically quite reliable. Further, because the simulation can be viewed from different angles, can be slowed down or speeded up, and zoomed-in as to improve perceptual access to salient parts of the folding process, it is fair to say that information access is highly reliable. Because programming the simulation can take hours to days and analysing the simulation usually takes many hours as well, the durability of the coupling is long. Simulations are usually analysed a number of times, so the relationship is not a one-off or permanent, but a repeated relationship is established.

Trust in the correctness and truth-value of the simulation is generally high, but this depends on the kind of simulation. The more is known about a particular protein, the more trust a modeller puts into its simulation. For example, when the tertiary structure is known via crystallography but not its folding process and when a simulation results in a similar or identical tertiary structure, then there is little reason to distrust the correctness of the simulation. But when simulating the folding of a protein whose tertiary structure is unknown, then a modeller may be more cautious in accepting the accuracy of the simulation. Note that accuracy here means degree of isomorphism between the simulation and the actual folding process. Simulations may of course be wrong in that the simulation is not isomorphic to the actual folding process. For

example, in some cases, the tertiary structure is known but not its folding process. When a simulation results in a structure that is not similar to its known tertiary structure, then modellers have good reason to distrust the simulation. If a high degree of trust is established, is it always after a modeller has consciously evaluated the simulation, so trust in case of simulations always concerns explicit trust.

The degree of procedural transparency depends on the degree of understanding of the software program, which may take some time and experience to develop. Using simulation software requires knowledge about both molecular biology and computer programming and the degree of transparency depends on this knowledge. For a novice, transparency is typically low, but for an experienced modeller it is high, with various degrees of transparency in between. The informational transparency is most likely high, as simulations of protein folding are highly isomorphic to the actual folding process. Even for complete novices, the basics of a simulation are relatively easy to understand, partly because simulations are approximations, simplifications, and idealisations of their target systems, which makes them easier to understand and easier to study.

Whether simulations are individualized depends on the research goals of the modeller. In some cases, simulations are performed by individual researchers working on their specific projects, in which case the simulation is highly individualized. In other cases, simulations are performed by a small team of researchers, in which case they are less individualized. But, either way, the outcomes of simulations are quite often published and are thus aimed at a larger audience and are performed to contribute to a better general understanding of molecular biology. In this sense, simulations are not individualized, but part of a publically accessible body of knowledge. Also, the simulations may sometimes be individualized but the computer software with which they are made is interchangeable, at least for those who master the program.

Computer simulations of protein folding and protein structure have most likely transformed the representational capacities of molecular biologists. Visualizing and perceiving how an unobservable such as a protein behaves in vivo, has changed how molecular biologists and others think about proteins. If a modeller spends many hours analysing simulations of protein folding, then it is likely that she will have internalized at least some parts of the simulation, but it is difficult to quantify how much computer simulations have transformed the representational capacities of their users. Also,

before computer simulations there were physical models of protein structure that most likely also transformed the way biologists think about such structures. But simulations are different as they are both structurally and sequentially isomorphic, whereas physical models are merely structurally isomorphic. This combined structural and sequential isomorphism has probably a stronger transformation effect on the modeller.

More generally, scientific models are representational systems that the brain is likely to absorb relatively easily. Particularly models that exhibit a high structural isomorphism to their targets are easy to internalize. For example, models of the solar system, the anatomy of the human body, atomic structures, plate tectonics, and many other models have changed the way we think about the physical world, not only for scientists but also for the general public, as some of these models have found their way into popular culture. This is because models make complicated processes or structures relatively easy to understand, as their isomorphic format is easy to interpret and internalize.

In sum, computer simulations of protein folding score relatively high on most dimensions. They score medium on speed of information flow, and durability. They score high on reliability and trust and after some training they also score high on procedural and informational transparency. They may or may not be individualized, depending on each case, and it is difficult to quantify how much they have transformed the representational capacities of their users, but it is safe to say that there is some degree of transformation. But, most importantly, there is a reciprocal information flow which means there is a significant degree of integration between agent and simulation. Thus, given how computer simulations of protein folding score on the dimensions, it populates a higher region in the multidimensional space and so integration between researcher and simulation is quite dense.

5.1.2 pH-Meters

A common procedure in molecular biology is to measure the acidity or alkalinity of a liquid, i.e., its pH. Obtaining this information is important as the pH determines some of the relevant properties of liquids such as, e.g., the amount of chemicals that will dissolve in it. In order to obtain the pH value, a series of nested actions needs to be performed: activating the device, dipping the measuring probe in the liquid, and interpreting its reading. All this is typically done in a few seconds, resulting in a pH value that is measured to two decimals, say, 5.68. When measuring pH, information

flow is one-way, i.e., from artifact to researcher. Information is not offloaded onto the device by its user, but it is possible to change the reading by altering the target system, for example by adding acidic or alkaline chemicals. Given that molecular biologists are trained to use pH-meters and have used such devices countless times, the information resulting from the epistemic actions is interpreted very quickly. So the speed of information flow is high.

Access to pH-meters is highly reliable, as there are usually a number of pH-meters available in any MBL. Moreover, pH-meters are regularly calibrated to ensure accurate readings, which increases the reliability of the information. Measuring pH takes only a few seconds, so the durability of each coupling is rather short. However, it is such a common procedure that the relation to pH-meters is often repeated and thus a long-term, permanent relationship is established. Trust in the accuracy of the reading is high, since pH-meters are regularly calibrated. Although there may be cases when the reading is unusual which may prompt suspicion or distrust, typically a researcher implicitly trusts the information it provides, i.e., she does not consciously evaluate it. For experienced researchers, the procedural and informational transparency are high, because they are trained in using pH-meters and have used them countless of times. The process is proceduralized and they do not have to consciously think about how to interact with the device or how to interpret its reading. Novices who lack sufficient training and experience may need more conscious thought to use the device and may experience it as less procedurally and informationally transparent.

pH-meters are not individualized, they are highly interchangeable and used by all members of the laboratory. In the previous chapter, I argued that in some cases cognitive functions are multiple realizable, i.e., different physical structures may perform identical functions. To a certain extent this is true for pH-meters, as one may also measure pH with litmus paper. Although litmus paper is much less accurate and has a very different physical structure as compared to electronic pH-meters, it does provide similar information about the target system. Finally, pH-meters may not have transformed the representational capacities of the researchers' brains in the same way as scientific models have, but they have transformed how we think about acidity and alkalinity. They have also transformed many cognitive practices in MBLs, because being able to accurately measure pH is important for many experimental procedures. The representational state of the device is dynamic and will change when the target system

changes, so it is not a static artifact. However, in terms of Nersessian's notion of cognitive partnership, which is characterized by an evolving and dynamic relation between agent and artifact, pH-meters do not qualify as artifacts with which a cognitive partnership is established, because their functional properties do not change over time.

In sum, pH-meters score high on most dimensions. They score high on all dimensions except on individualization and transformation on which they score low, and, importantly, the information flow is one-way. Given how it scores on the above dimensions, this situated cognitive system populates a region somewhere in the lower middle regions of the multidimensional space. But, if the information flow was two-way or reciprocal, then it would have scored significantly higher and the system would be integrated more densely. More generally, one can say that indices do not afford two-way or reciprocal information flow, because they have a direct causal connection to their target systems. One can change their informational content only by altering the target system. So one can reason *about* the information an index provides and one may use that information to guide further epistemic or pragmatic action, but one cannot reason *with* indices in the same way as one can with icons and symbols. This is so because they lack representational malleability, i.e., the capacity to change their content in an ongoing cognitive task. It is therefore difficult for an index and its user to be deeply integrated into a larger system. Recall that one of the reasons for taxonomizing cognitive artifacts was to examine whether artifacts in different categories, i.e. with different informational properties, have different effects on cognition. Because indices lack the capacity for two-way or reciprocal information flow, they also lack the capacity for deep integration.

5.1.3 The Laboratory Notebook

pH-meters have only one cognitive function (to measure pH) and exhibit only one type of information (a numerical value of pH). Laboratory notebooks, by contrast, are much more complicated as they serve a number of cognitive functions (mainly related to memory and reasoning) and exhibit various types of information (language, equations, tables, numbers, etc.). The *cognitive* role of laboratory notebooks has not received much attention, neither in the situated cognition literature nor in the philosophy of science

literature[54]. Lab notebooks are used for various cognitive tasks and serve a number of cognitive functions[55]:

- When performing an experiment, it is very hard to remember all the steps as most experiments are rather complex and involve countless incremental steps. Therefore, lab notebooks always contain a section on methodology, which is typically written by the researcher, describing all the steps that need to be done in order to perform the experiment. In this role, the notebook functions as a to-do list, so that the researcher knows what to do and what has been done at any point in the experiment.
- Observations made during the experiment are written in the notebook. The authors of a biology textbook write: "Record all observations as you make them. Do not trust to memory even for a minute" (Thompson & Thompson 2012, p. 5). This quote indicates that researchers are aware that human memory is limited. In this role, the notebook functions as a long-term external information-storage system, thereby complementing the shortcomings of internal storage systems.
- Experimental outcomes are written in the notebook, sometimes as linguistic descriptions but more often as tables, graphs, or other diagrams (Kanare 1985). In this way, researchers later know what they have done during an experiment and what the outcome of the experiment was. These experimental outcomes are often discussed during lab meetings where notebooks plays a crucial role as memory aids.
- In addition to these predominantly memory aiding functions, notebooks are also used to solves equations, perform calculations, or draw graphs, tables, or other diagrams. In these roles, the notebook serves more as a facilitator of ongoing reasoning processes.

---

[54] An exception is Richard Yeo (2008) who studied the cognitive role of notebooks and note-taking in scientific practice in 17th century England. Notebooks have also been used to reconstruct historical events. Faraday's notebooks, for example, have been used for these purposes (e.g. Tweney 1991). Further, the cognitive role of notebooks has been addressed by Clark and Chalmers (1998), but not for healthy agents and not in the context of a laboratory.

[55] Perhaps one would expect that lab notebooks would now be predominantly electronic, but a recent review suggests that the majority of lab notebooks is still paper-based (Rubacha, Rattan & Hosselet 2011).

Depending on the task for which it is used, information flow between notebook and user is two-way or reciprocal. When an experimental procedure is first written in the notebook and then later used to help organise the experiment, information flow is two-way because there are two steps involved: offloading and intake. But when it is used to perform calculations, solve chemical equations, or draw graphs or tables, then a reciprocal system is established. This is so because performing these tasks typically involves many cycles of offloading and intake and each cycle depends on the outcome of the previous one. Each step in the overall cognitive task thus builds and depends on and the previous step and are therefore interdependent.

Offloading speed is relatively quick, since writing is a fairly quick method to offload information onto the environment. Informational uptake is also fairly quick as the information is written by the user herself, so it is her own handwriting and the information is typically structured towards easy intake. Thus, generally, the speed of information flow is medium to fast, but there may be exceptions where a researcher has to solve a difficult calculation or equation which may take some time to think about.

Access to the notebook is highly reliable, as it is standard procedure to always have it around when performing experiments. Some researchers make photocopies of their lab notebooks so that they have a backup of it, which increases reliability. The information itself is, in most cases, also highly reliable. But there may, of course, be cases in which mistakes are made when performing calculations, solve equations, and so on, in which case the information is less reliable or unreliable. Given that the notebook is used many times throughout the day, a repeated relationship is established. The notebook is as important for a molecular biologist as a hammer for a carpenter.

Most of the information in the notebook is implicitly trusted, i.e., not consciously evaluated, because it is written by the researcher herself, but also because some of the information is standardized. Most experimental procedures, for example, are standardized so that they need little evaluation when used to guide the experiment. But when interpreting experimental outcomes, more evaluation is needed and the data may not be trusted by default. After conscious evaluation, a researcher may either explicitly trust the data or distrust it.

Interacting with notebooks is rather straightforward, so the procedural transparency is high. The information in it (e.g., sentences, calculations, equations) is written by the agent herself, so it is easy to interpret and process. Although there may be cases in which a researcher has quickly written some observations in the notebook that are hard to interpret or understand at some later point. But generally the informational transparency is high. The notebook is highly individualized and is deeply engrained in many of the cognitive practices of the individual user. The information in the notebook is meant to perform her specific experiments and the observations she writes in it result from her experiments. So it is as individualized as a cognitive artifact can be. Researchers also deeply depend on it. If the notebook got lost, a researcher would have trouble performing experiments and the experimental data would be lost.

Finally, the notebook most likely does not transform the representational capacities of their users. Only interacting with representational systems such as language, mathematics, and perhaps some notational systems in science like chemical equations or models will significantly transform the representational capacities of embodied brains. Single artifacts most likely do not transform the brain in the same way representational systems do. However, artifacts do (sometimes quite dramatically) transform our cognitive practices. Lab notebooks, for example, have significantly transformed many of the cognitive practices in laboratories (see, e.g., Holmes, Renn & Rheinberger 2003). In sum, lab notebooks score high on all dimensions, except speed of information flow and transformation. They furthermore exhibit a two-way and reciprocal information flow. Given this score, notebooks populate a higher region in the multidimensional space and, therefore, notebooks and their users are deeply integrated into a distributed cognitive system.

5.2 Ecological Artifacts

5.2.1 Organized Workspaces

Molecular biologists organise their work-environment such that the location of the artifacts used in performing their experiments facilitates the cognitive tasks they are performing. This "intelligent use of space" (Kirsh 1995) reduces the cognitive load in perceptual and memory tasks, in that way complementing their cognitive processes. When preparing the experiment by intentionally putting artifacts in particular locations on one's workbench, task-relevant information is encoded into the artifacts and their locations. So information is first offloaded onto the environment by intentionally

putting artifacts at certain locations and then taken onboard at some later point. A two-way information flow system is thus established. It takes some time to prepare the experimental setup and performing the experiment usually takes a fair bit of time as well. So the speed with which information flows in this two-way system is medium.

Access to the information is highly reliable, as the information is always there when it is needed, i.e., when performing the experiment one is sitting at one's organised workplace which is central in one's visual field. Performing an experiment can take thirty minutes or several days. So depending on each experiment, the duration of the coupling is medium to very long. Organised workplaces are repeatedly created and in that sense a permanent relationship is established with such ecological artifacts. However, given that each experiment may require different instruments, set-ups, and procedures, the ecological artifacts that are created for each experiment will have different structures and will have different informational properties. So there is no permanent relation established to specific ecological artifacts, unless an experiments is performed many times, but typically new ecological information is created for each experiment.

The amount of trust in the correctness of the information is generally high, as the researcher herself has placed the artifacts in their correct location. When performing the experiment, a researcher most likely will not consciously evaluate whether the location of artifacts indicates the correct order of steps in the experiment, but will implicitly trust it is correct. The procedural transparency is high, as the researcher herself has put the artifacts in certain locations and she knows how to extract the task-relevant information. Similarly, the informational transparency is high, too, for the same reason. Organized workplaces are highly individualized, as it concerns the specific experiments of individual researchers. Finally, organised workplaces most likely do not transform the representational capacities of their creators, because they do no concern external representations.

In sum, organized workspaces score low on transformation, medium on speed of information flow, the durability depends on the kind and complexity of the experiment, and they score high on reliability, trust, procedural and informational transparency, and individualization. Given how organised workspaces score on the dimensions, they

populate an upper middle region of the multidimensional space and are thus not deeply integrated.

5.3 Discussion of Dimensions

Which dimensions are more important for conceptualizing the degree of integration? Information flow is one of the most important dimensions. Situated cognitive systems in which there is a one-way information flow, as is the case with indices like pH-meters, but also with icons (e.g. maps) and symbols (e.g. timetables, textbooks) are not densely integrated. Situated cognitive systems with two-way or reciprocal information flow are integrated much deeper, because the user has intentionally created the information.

Speed of information flow seems to be less important, although this depends on the task. If, for example, a pilot in a cockpit quickly needs some unit of information, then speed of information flow is essential to form an effectively integrated cognitive system. But for most cognitive tasks the speed with which the information flows is not essential for the degree of integration. A simulation of protein folding which is run very slowly can still be deeply integrated with the cognitive system of its user. In fact, in some cases a slow information flow may even increase integration, as it allows close inspection of salient properties of the folding process.

Reliability is also quite important because if the artifact is not available, there can be no integration at all. In this sense, reliability is the most important dimension. If computer simulations, pH-meters, notebooks, organised workplaces, and other cognitive artifacts are not available when needed, there can be no integration at all. Durability seems somewhat less important, as a one-off such as a shopping-list can still be integrated fairly deeply into the cognitive system of its user.

Trust in information is important. Ideally we implicitly trust the information, in which case we do not think about its truth-value. In such cases, the information can be used quickly and fluently to guide further thought or action. If we distrust information, we typically do not or cannot use it to guide our thoughts or actions. So the more we trust the information, the more fluently the situated system works.

If the artifact is not procedurally transparency, then integration becomes difficult. For example, if an agent does not know how to interact with computers or pH-meters due

to lack of training and experience, it is impossible to actually use the artifact. Likewise, if the artifact is not informationally transparency (i.e. when its user does not understand the information), then integration is close to impossible. If the meaning of the information is not understood by the agent, then it cannot be used to perform some cognitive task.

The amount of individualization is also somewhat less important. Generally, the more an artifact is individualized, the more it is used and the deeper it is integrated. However, interchangeable artifacts such as word-processors and using pen and paper to perform calculations can still be deeply integrated with the cognitive system of their users, due to the reciprocal information flow.

Transformation, in a sense, is more the foundation for integration. If certain representational systems are not sufficiently internalised, we cannot use them externally. So, for example, if we have not learned to use language or mathematics, we cannot actually use those systems externally. So in this sense, transformation is more a condition for integration. Thus, information flow, reliability, trust, procedural and information transparency are important dimensions, whereas speed of information flow, durability, and individualisation and transformation are somewhat less important.

6. Concluding Summary

In this chapter, I first reviewed current classifications of scientific instruments and argued that a classification focussing exclusively on cognitive artifacts in MBLs is a valuable addition to current classifications, which focus on instruments in general and overlook a number of important classes of cognitive artifacts such as symbolic artifacts and ecological artifacts. I then grouped a number of cognitive artifacts in terms of the taxonomy developed in chapter 4, thereby creating structure and systematicity. Having classified some of the artifacts in MBLs, I reviewed how some theorists have used a distributed cognition approach to better understand aspects of scientific practice. I argued that although these approaches are perceptive and insightful, they would benefit from a more detailed look at the micro-interactions between researcher and artifact, as these determine how extended/distributed/integrated a situated cognitive system is. Finally, I performed four case studies and conceptualized how and how deeply computer

simulations of protein folding, pH-meters, lab notebooks, organized workspaces and their users are integrated into larger distributed systems.

7

Conclusions and Future Research

1. Conclusions

The main goal of this thesis was to better understand the variety of situated cognitive systems consisting of embodied agents and cognitive artifacts, and to conceptualize how such artifacts and their users are integrated into systems that perform cognitive tasks (in scientific practice). To this end, I started by reviewing and discussing several characterizations and classifications of artifacts that aid their users in performing cognitive tasks. I discussed the notions of exograms (Donald), cognitive artifacts (Norman, Brey, Hutchins, Nersessian), epistemic tools (Sterelny), and cognitive technology (Clark) and pointed out that the defining property of such artifacts is their functional role in performing a cognitive task. More specifically, their function is to provide task-relevant information, in that way making certain cognitive tasks easier, faster, more reliable, or possible at all. In order to better understand the distinctiveness of artifacts with cognitive functions, I contrasted cognitive functions with pragmatic functions by building on and further developing the notion of an embodied tool, which typically (though not necessarily) lacks cognitive functions.

I also discussed the notions of mental artifacts (Norman), internal artifacts (Hutchins), and internalized cognitive artifacts (Sutton). On the basis of those notions, I developed

the concept of internalized information and distinguished between two broad categories of cognitive techniques, those that substitute cognitive artifacts (e.g. mentally visualizing and manipulating abacus beads to perform a calculation) and those that concern interactive functional skills that allow us to exploit our environment for cognitive purposes (e.g. our learned ability to interpret language). In addition to human-made or artificial scaffolds - i.e., artifacts and techniques - we also deploy cognitive naturefacts (e.g. navigating with the aid of stars) and other people (e.g. long married couples that complement each other's memory systems) to help us perform our cognitive tasks. Thus the preliminary taxonomy of components of situated cognitive systems, as developed in chapter 2, consists of artificial, natural, and social scaffolds.

In chapters 3 and 4, I further conceptualized some of the metaphysical properties of cognitive artifacts. I began by identifying the target domain by defining this class of artifacts as a functional kind. More specifically, this functional kind includes cognitive artifacts with both proper and system functions. Those with proper functions have a history of cultural selection, whereas those with system functions are improvised uses of initially non-cognitive artifacts. Next, by drawing on artifact categorization in archaeology, a taxonomy of cognitive artifacts was developed. In developing this taxonomy, an information-centered approach was taken, i.e., I took as my point of departure the specific informational properties of cognitive artifacts and then taxonomized them on the basis of those properties. Current categorizations focus on representational artifacts and thus neglect non-representational or ecological artifacts. They also tend to focus on proper cognitive artifacts and often overlook system cognitive artifacts (Kirsh being a notable exception). These categorizations therefore have a significantly smaller target domain. Moreover, all current categorizations are cognition-centered in that they start with human cognition and then work their way towards the (proper) representational artifacts that agents deploy to aid their cognition and realize their cognitive purposes. My information-centered approach is a valuable alternative to these approaches, as it results in a much richer and detailed taxonomy.

In this taxonomy, three levels or taxa are distinguished, those of family, genus, and species. The family includes all cognitive artifacts (i.e. proper, system, representational, and ecological) without further specifying functional or informational properties. On the second level in the taxonomy, I distinguished between two genera, those of representational and ecological cognitive artifacts. On the third level, these two genera

are further divided into species. In case of representational artifacts, those species are: iconic, indexical, and symbolic. In case of ecological artifacts, those species are: spatial and dynamic. Within species, I identified type cognitive artifacts. The categories in the taxonomy are not mutually exclusive, but one particular informational property is usually predominant.

In addition to taxonomizing cognitive artifacts, I also conceptualized one of their fundamental metaphysical properties, i.e., their structure-function relations. In case of representational artifacts, I particularly looked at the relation between the artifact's physical structure and its representational structure. I argued that representational structures can be (1) carried by or added onto a physical structure, e.g., a map printed on paper; (2) statically or dynamically constituted by physical structures, e.g., a scale model or thermometer; and (3) computed by a physical structure, e.g. slide ruler or digital computer. Ecological artifacts lack representational properties and therefore have different structure-function relations. In such artifacts, information can be encoded as follows: (1) physical structures can be intentionally placed at a particular location, thereby creating a hybrid of a physical-spatial structure, from which task-relevant information emerges; or (2) information-carrying structures can be manipulated such that new information emerges from their new configurations.

In chapter 5, I first briefly compared parity-based and complementarity-based extended cognition theory. The parity principle stresses functional parity between internal and external states and processes, whereas the complementarity principle stresses complementarity between the internal and the external. By stressing functional isomorphism, the parity principle downplays differences between internal and external states and processes, in that way implying that the nature and properties of cognitive artifacts and their impact on our brains and behaviour do not really matter. Additionally, it also downplays individual differences between humans and how they interact with cognitive artifacts. On a complementarity view, biological and artifactual elements in situated cognitive systems need not exhibit similar properties or functionalities. Rather, embodied agents deploy the different functional and informational properties of cognitive artifacts to complement their onboard cognitive capacities.

To further develop a complementarity view and to better understand the functional integration of agents and artifacts, I developed a multidimensional framework for conceptualizing the degree of integration. This framework consists of the following dimensions: epistemic action and information flow, speed of information flow, reliability, durability, trust, procedural transparency, informational transparency, individualization, and transformation. The proposed dimensions are all matters of degree and jointly they constitute a multidimensional space in which situated cognitive systems can be located. Importantly, these dimensions provide a new perspective on the conditions for cognitive extension and cognitive distribution. However, they are not meant to provide a set of necessary and sufficient conditions, but to provide a toolbox for investigating the degree and nature of the integration of agent and artifact into "new systemic wholes". The higher a situated system scores on the proposed dimensions, the more functional integration occurs, and the more tightly coupled the system is. How these dimensions can be utilized for empirical research is outlined below in section 2.

In the final chapter, I first reviewed current classifications of scientific instruments and argued that a classification focussing exclusively on cognitive artifacts in molecular biology labs is a valuable addition to current classifications, which focus on instruments in general and overlook a number of important classes of cognitive artifacts such as symbolic artifacts and ecological artifacts. I then grouped a number of cognitive artifacts in terms of the taxonomy developed in chapter 4. Having classified some of the artifacts in MBLs, I reviewed how some theorists have used a distributed cognition approach to better understand aspects of scientific practice. I argued that although these approaches are perceptive, they would benefit from a more detailed look at the micro-interactions between researcher and artifact, as these determine how extended/distributed/integrated a situated cognitive system is. Lastly, I performed four case studies and conceptualized how and how deeply computer simulations of protein folding, pH-meters, lab notebooks, organized workspaces and their users are integrated into distributed cognitive systems.

2. Future Research

I conclude by pointing out some suggestions for further conceptual and empirical research by focussing on enlarging the explanandum, how this thesis may benefit from and guide further empirical research, and some normative issues.

2.1 Enlarging the Explanandum

This thesis has focused mainly on small-scale situated systems, consisting of individual agents and cognitive artifacts. In some cases, however, small-scale systems are comprised of individual agents and other external resources such as cognitive naturefacts or other people like, for example, navigating with the aid of stars or long married couples. These have been briefly presented in chapter 2, but finer-grained taxonomies of these components, outlining their functions and informational properties, would be helpful for situated cognition theory. Likewise, the multidimensional framework was designed to conceptualize the integration between individual agents and artifacts, but I see no reason why these dimensions cannot be applied to systems consisting of other components. Thus, when analysing socially distributed systems such as long married couples (Sutton et al 2010) or sports teams (Williamson & Cox 2013), the multidimensional framework would be useful for conceptualizing the degree of integration.

2.2 Empirical Research

The conceptual research performed in this thesis could be tested by and form the basis for empirical research. Particularly, the multidimensional framework could be used to guide and inform empirical research concerning agent-artifact interactions in particular contexts such as, for example, a scientific laboratory. However, this is not an easy task because different dimensions need different empirical methods, drawing on experimental psychology, cognitive anthropology, and neuroscience. This pluralism in methodology makes it difficult empirically to study every case of agent-artifact interaction in terms of all the dimensions. In most cases, however, most of the dimensions can be empirically studied. An additional challenge is specifying quantifiable values for each dimension. In chapter 6, I used three values: low, medium, and high. These were helpful heuristically, but more precise descriptions of these three values for each dimension would strengthen the framework.

Before I discuss the empirical methods appropriate for studying these dimensions, I point out two broad approaches to investigating integration in situated cognitive systems. Depending on the kind of situated system, one may study integration either in an artificial or natural setting, i.e., in a laboratory or "in the wild". Some situated systems are best studied in a psychology laboratory. For example, if one is interested in how agents play Tetris and how the outsourcing of the rotation of zoids is integrated into the onboard cognitive system, a laboratory study is most suitable (see Kirsh & Maglio 1994). This is so because a laboratory provides a controlled environment and observational tools such as camera's, eye-tracking devices, and other recording devices. And, moreover, playing Tetris is arguably not relevantly different in a laboratory setting, as compared to a naturalistic setting. Likewise, if one wants to study how agents assemble a piece of furniture with the aid of a manual and/or by placing the parts such that they correspond to the most logical order of actions, a laboratory study would be appropriate, because it is much easier to record and analyse, and is not relevantly different from a naturalistic setting.

Other situated systems, however, are impossible to study in a laboratory. If one is interested in, for example, the navigation practices of the Micronesian people (Hutchins 1984, 1995) or how scientists use cognitive artifacts in their research practices (Nersessian 2005, 2006), an "in the wild" approach is necessary. That is to say, such situated systems can only be meaningfully studied in a naturalistic setting. Therefore, some systems, regardless of the dimensions one is studying, are inherently impossible to study in a psychology laboratory, which has consequences for the dimensions one can study. In some cases, situated systems may be studied in both an artificial and naturalistic setting. Pilots in a cockpit, for instance, can be studied both in an actual cockpit (see Hutchins 1995b) and a flight simulator (see Hutchins & Klausen 1996), although there could be relevant differences such as safety issues which may influence how pilots interact with the cockpit.

These two approaches come with limitations and advantages. Artificial settings are limited to a relatively small amount of situated systems such as different forms of Human-Computer Interaction, diagrammatic reasoning, and other relatively straightforward problem-solving activities involving cognitive artifacts. The advantages of artificial settings are a highly controlled environment and recording methods. Naturalistic settings are much less limited in the amount of systems one can study, but

may suffer from a lack of observation and recording methods. It is, for example, difficult to consistently and diachronically record how a team of mountaineers use navigational artifacts such as maps, compasses, or GPS-based navigation devices to navigate on Mount Everest.

Having pointed out two empirical approaches one may take to investigate agent-artifact integration, I now briefly explain how each dimension can be empirically investigated. All the dimensions can, at least in principle, be investigated by observing how agents interact with cognitive artifacts, either in artificial or naturalistic settings. Mere observations may, however, not always yield satisfactory results and so additional qualitative methods such as (semi-structured) interviews or questionnaires can supplement the information obtained from observations. In some cases, neuroimaging studies may provide additional information. For example, if one is interested in cognitive transformation, fMRI studies can show how the brain's structure changes when skills are learned and when external representational systems are soaked-up (see e.g. Dehaene et all 1999).

The dimensions related to information trajectories, i.e., epistemic action and direction of information flow and speed of information flow can be investigated by observations. It is relatively straightforward to obtain how information trajectories are established in situated systems. The epistemic actions of agents and the informational properties and functions of the artifacts are overt and can be observed. Sometimes additional methods are needed. For example, epistemic actions such as saccades to external information can be recorded by using eye-tracking technology. In some cases, the amount of information that flows between the components of a situated system may be too large to observe and record. Writing an academic paper can take months with countless cycles of reciprocal information flow, which is in practice very hard to observe due to practical limitations in analysis methods. Only relatively simple information trajectories or short episodes of highly durable situated cognitive systems can in practice be observed.

When navigating with the aid of a map, the epistemic actions (e.g., orientating the map such that it is perceivable) as well as the informational properties of the map can be observed, but it is much more difficult to observe what information an agent actually takes onboard. There are basically two (complementary) ways in which one may obtain what information an agent actually takes onboard. First, by observing behaviour. For

instance, after looking at a map, an agent walks into a certain direction, based on that behaviour one may infer that the agent took certain task-relevant information onboard. Second, by interviewing the agent. After the cognitive task is completed, a researcher can ask in a (semi-structured) interview or perhaps with questionaries, what information the agent (thinks he or she) has actually taken onboard and used for completing the task. Speed of information flow can be obtained by measuring the duration of the cognitive task as well as the amount of information that is used. However, the amount of task-relevant information may not always be quantifiable and depends on the kind of informational system.

Reliable access and durability only make sense to observe in a naturalistic setting. How durable an artifact is and how often an agent makes use of it can be quite easily observed. So, for example, an ethnographer observing how scientists use their instruments in a scientific laboratory, can quite easily observe how often an artifact is used. Trust may be inferred from observing behaviour. For instance, if, after looking at a map, an agent walks into a certain direction, one may infer that the agent trusted the information provided by the map. Whether trust is implicit (accepted without conscious evaluation) or explicit (accepted after conscious evaluation) can only be obtained by means of interviews.

The degree of procedural and informational transparency can be investigated by observing how skilfully and efficiently an agent uses a cognitive artifact. Likewise, how skilfully and efficiently an agent interprets information can also be observed. Given that procedural and informational transparency develop and (ideally) increase over time, they need to be studied diachronically. That is, a series of observations have to be done over a certain period of time to observe how an agent's skills improve. If information is non-transparent or not sufficiently transparent, an agent cannot use it to complete the cognitive task. So, if an agent is confronted with a certain cognitive task and is unable to complete it due to a lack of transparency, this is observable. How artifacts are individualized and how the system is transformed is relatively straightforward to observe, as epistemic actions and the informational properties of cognitive artifacts are overt. Given that these dimensions occur over time, they need to be studied diachronically. This may be done by longitudinal studies or by doing a series of snapshots at distinct developmental stages.

It may not always be possible to study all these dimensions in every possible case. It is, for example, difficult (though in principle not impossible) to test how much neural transformation occurs when someone consistently leaves her car keys at a certain spot in her apartment, as the effect may be too insignificant to be measureable. In most cases, however, most of the dimensions can be empirically studied. To conclude this subsection, in order to empirically investigate the integration between agent and external resource, I suggest - depending on each individual case - a combination of ethnography, questionaries and interviews, eye-tracking technology, and neuroimaging to jointly research the dimensions.

2.3 Normative Questions

In addition to enlarging the explanandum and empirical work, the research in this thesis (and situated cognition theory more generally) may benefit from a normative analysis. There are very few explicit connections between situated cognition theory and moral philosophy, exceptions are Neil Levy (2007), Mason Cash (2010) and Zoe Drayson & Andy Clark (in press). To further build such connections, a number of normative issues might be addressed. If, under certain conditions, external artifacts and structures are literally part of the mind, then intervening with these objects is intervening with the mind (Levy 2007). Following this line of reasoning, certain cognitive artifacts and other external structures have a special moral status, because rather than being mere objects, they are part of a human cognitive system. When human thought is externalized and publically accessible, moral values like informational privacy and informational security become relevant for a full conceptualization of extended/distributed cognitive systems. Here I think situated cognition theory could draw on the ethics of technology (e.g. Floridi & Sanders 2004; Kroes & Verbeek 2014) to conceptualize the moral dimensions of cognitive artifacts, thereby further building intradisciplinary bridges between different fields in philosophy.

A second issue concerns the desirability of the effects of cognitive artifacts on our brains and onboard cognitive system. In ontogeny, language, mathematics, and some other representational systems are soaked up by the brain, and we learn how to use and manipulate these systems to perform cognitive tasks. This hardly seems problematic from a moral perspective as these cognitive capacities largely define us as human beings and have caused substantial progress for humans and society at large. However, an overreliance on external systems may cause a diminishing of some of our cognitive

capacities as some of these representational systems may not be (sufficiently) internalized and may either not transform our brain at all, or transform our brain in undesirable ways. Consistently outsourcing information-storage and information-processing functions to artifacts that are to varying degrees integrated into larger systems may result in a loss of the outsourced function or capacity. An often heard critique is that cognitive artifacts do not make us smart but make us stupid (see, e.g., Carr 2011). The idea is that if artifacts do all the cognitive work for us, we never learn to do it ourselves, or, when we have learned to do it ourselves, we might lose that capacity due to a lack of practice.

This worry echoes Socrates' critique of written language. Socrates argued that written language would erode memory as it allows us to store information in the environment, rather than in the brain, thereby making us cognitively lazy. Although Socrates may have a point[56], very few people today would think that the development of written language is morally or culturally undesirable. A contemporary version of Socrates' worry is that our digital cognitive artifacts allow reliable information access and, as a result, we never (learn to) memorize information, i.e., we learn information pathways rather than information itself. One way to look at this issue is to see it as a trade-off between losing something and gaining something. If we gain more than we lose, then the trade-off is acceptable. Note that this may not be easy as it is difficult to precisely determine what we have lost and what we have gained from any given (category of) cognitive artifacts or informational systems. This consequentialist view may not be accepted by those who think that having certain cognitive skills has intrinsic value, and whose value should not be measured in relation to what we might gain. Non-consequentialists would argue that those skills are valuable in themselves and should be cultivated, regardless of the benefits cognitive artifacts may bring.

I think it is probably too early to fully evaluate the cognitive and cultural consequences of these developments, but I do not think that these developments are fundamentally different from previous ones. From an evolutionary perspective, these developments do not seem too problematic. If we are "natural-born cyborgs" (Clark 2003), then it is in our basic human nature to incorporate tools and artifacts into our bodily, perceptual,

[56] Somewhat ironically, if Plato did not write down Socrates' worries in his *Phaedrus*, we would probably not know about it today.

and cognitive systems, in that way complementing those systems. As Donald (1991) has argued, during the evolution of our cognition, we have developed (external) representational systems that have had an evolutionary impact on our cognition. This co-evolutionary process between embodied brains and external information-storage systems has been occurring at least since the invention of the first external information-storage systems and so the effect of digital artifacts on our cognition is not different in kind than those of previous artifacts. I am not saying that because this co-evolutionary process has been occurring for a very long period, it is natural and therefore acceptable. This would be committing the naturalistic fallacy. What I am saying is that we need to carefully investigate and evaluate these developments and be aware of their possible consequences.

References

Achterhuis, H. (ed.) (2001). *American Philosophy of Technology: The Empirical Turn*. Indiana University Press.

Ackerman, R. (1985). *Data, Instruments, and Theory: A Dialectical Approach to Understanding Science*. Princeton University Press.

Adams, W.Y. & Adams, E.W. (1991). *Archaeological Typology and Practical Reality: A Dialectical Approach to Artifact Classification and Sorting*. Cambridge University Press.

Adams, F. & Aizawa, K. (2001). The Bounds of Cognition. *Philosophical Psychology,* 14, 43-64.

Adams, F. & Aizawa, K. (2008). *The Bounds of Cognition*. Wiley-Blackwell.

Adams, Z. & Maher, C. (2012). Cognitive Spread: Under what Conditions does the Mind Extend Beyond the Body? *European Journal of Philosophy*. DOI: 10.1111/j.1468-0378.2012.00543.x.

Alač, M. & Hutchins, M. (2004). I See What You are Saying: Action as Cognition in fMRI Brain Mapping Practice. *Journal of Cognition and Culture,* 4(3), 629-661.

Anderson, M.L. (2003). Embodied Cognition: A Field Guide. *Artificial Intelligence,* 149, 91–130.

Atkin, A. (2008). Icon, Index, and Symbol. In P. Hogan (ed.), *The Cambridge Encyclopaedia of Language Sciences* (pp. 367-368). Cambridge University Press.

Atkin, A. (2010). Pierce's Theory of Signs. In *Stanford Encyclopedia of Philosophy* (Fall 2012 Edition), E.N. Zalta (ed.), http://plato.stanford.edu/entries/peirce-semiotics/.

Baird, D. (2003). Thing Knowledge: Outline of a Materialist Theory of Knowledge. In H. Radder (ed.), *The Philosophy of Scientific Experimentation* (pp. 39-67). University of Pittsburgh Press.

Baird, D. (2004). *Thing Knowledge: A Philosophy of Scientific Instruments*. University of California Press.

Bakker, L.R. (2004). The Ontology of Artifacts. *Philosophical Explorations,* 7(2), 99-111.

Barnier, A.J., Sutton, J., Harris, C.B., & Wilson, R.A. (2008). A Conceptual and Empirical Framework for the Social Distribution of Cognition: The Case of Memory. *Cognitive Systems Research*, 9(1), 33-51.

Beach, K. (1988). The Role of External Mnemonic Symbols in Acquiring an Occupation. In M.M. Gruneberg & R.N. Sykes (eds.), *Practical Aspects of Memory, Vol. 1* (pp. 342–346). Wiley.

Bechtel, W. & Abrahamsen, A. (2013). Roles of Diagrams in Computational Modelling of Mechanisms. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 1839-1844).

Becvar, L.A., Hollan, J., Hutchins, E. (2005). Hands as Molecules: Representational Gestures as Cognitive Artifacts for Developing Theory in a Scientific Laboratory. *Semiotica*, 156, 89-112.

Becvar, L.A., Hollan, J., Hutchins, E. (2007). Representational Gestures as Cognitive Artifacts for Developing Theory in a Scientific Laboratory. In C. Halverson & M. Ackerson (eds.), *Artifacts in Workplace Practice* (pp. 117-143). Kluwer Academic Publishers.

Boon, M. (2004). Technological Instruments in Scientific Experimentation. *International Studies in the Philosophy of Science,* 18(2-3), 221-230.

Boon, M. (2009). Instruments in Science and Technology. In J.K.B. Olsen, S.A. Pederson & V.F. Hendricks (eds.), *A Companion to the Philosophy of Technology* (pp. 78-83). Wiley-Blackwell.

Burks, A.W. (1949). Icon, Index, and Symbol. *Philosophy and Phenomenological Research,* 9(4), 673-689.

Brey, P. (2000). Technology and Embodiment in Ihde and Merleau-Ponty. In C. Mitcham (ed.), *Metaphysics, Epistemology, and Technology: Research in Philosophy and Technology, vol 19*. Elsevier/JAI Press.

Brey, P. (2003). Wetenschaps- en Techniekfilosofie. Does it take two to Tango? *Algemeen Nederlands Tijdschrift voor Wijsbegeerte*, 95(3), 215-20.

Brey, P. (2005). The Epistemology and Ontology of Human-Computer Interaction. *Minds and Machines,* 15(3-4), 383-398.

Brey, P. (2010). Philosophy of Technology after the Empirical Turn. *Techné,* 14(1), 36-48.

Brey, P. & Søraker, J.H. (2009). Philosophy of Computing and Information Technology. In A. Meijers (ed.), *Philosophy of Technology and Engineering Sciences* (pp. 1341-1408). Elsevier.

Brooks, R. (1999). *Cambrian Intelligence: The Early History of the New AI*. MIT Press.

Brooks, R. (2002). *Flesh and Machines: How Robots Will Change Us*. Pantheon.

Carr, N. (2011). *The Shallows: What the Internet is Doing to our Brain*. W.W. Norton & Company.

Carrara, M. & Vermaas, P. (2009). The Fine-Grained Metaphysics of Artifactual and Biological Functional Kinds. *Synthese,* 169(1), 125-143.

Carrara, M. & Mingardo, D. (2013). Artifact Categorization: Trends and Problems. *Review of Philosophy & Psychology,* 3(4), 351-373.

Cash, M. (2010). Extended Cognition, Personal Responsibility, and Relational Autonomy. *Phenomenology and the Cognitive Sciences*, 9(4), 645-671.

Cavalli-Sforza, L.L. & Feldman, M.W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press.

Charbonneau, M. (2013). The Cognitive Life of Mechanical Molecular Models. *Studies in History and Philosophy of Biological and Biomedical* Sciences, 44(4), 585–594.

Cheng, P. (2009). Probably Good Diagrams for Learning: Representational Epistemic Recodification for Probability Theory. *Topics in Cognitive Science*, 3, 475-498.

Cheon, H. (2013). Distributed Cognition in Scientific Context. *Journal for General Philosophy of Science*. DOI: 10.1007/s10838-013-9226-4.

Clancey, W.J. (1997). *Situated Cognition: On Human Knowledge and Computer Representations*. Cambridge University Press.

Clark, A. (1989). *Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing.* MIT Press.

Clark, A. (1997). *Being There: Putting Brain, Body and World back Together Again*. MIT Press.

Clark, A. (2001). *Mindware: An Introduction to the Philosophy of Cognitive Science*. Oxford University Press.

Clark, A. (2002). Towards a Science of the Bio-technological Mind. *International Journal of Cognition and Technology,* 1(1), 21-33.

Clark, A. (2003). *Natural Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford University Press.

Clark, A. (2007). Reinventing Ourselves: The Plasticity of Embodiment, Sensing, and Mind. *The Journal of Medicine and Philosophy,* 32(3), 263–282.

Clark, A. (2008a). Pressing the Flesh: A Tension in the Study of the Embodied, Embedded Mind? *Philosophy and Phenomenological Research,* 76(1), 37-59.

Clark, A. (2008b). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford University Press.

Clark, A. (2010). Memento's Revenge: Extending the Extended Mind. In R. Menary (ed.), *The Extended Mind* (pp. 43-66). MIT Press.

Clark, A., & Chalmers, D. (1998). The Extended Mind. *Analysis,* 58, 10-23.

Clowes, R.W. (2013). The Cognitive Integration of E-Memory. *Review of Philosophy and Psychology,* 4(1), 107-133.

Cummins, R. (1975). Functional Analysis. *The Journal of Philosophy,* 72(20), 741-764.

Cummins, R. (1983). *The Nature of Psychological Explanation*. MIT Press.

Cummins, R. (1996). *Representations, Targets, and Attitudes*. MIT Press.

Crilly, N. (2010). The Roles that Artifacts Play: Technical, Social and Aesthetic Functions. *Design Studies,* 31, 311-344.

Dascal, M. & Dror, I.E. (2005). The Impact of Cognitive Technologies: Towards a Pragmatic Approach. *Pragmatics and Cognition,* 13(3), 451-457.

De Cruz, H. (2008). An Extended Mind Perspective on Natural Number Representation. *Philosophical Psychology,* 21(4), 475-490.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of Mathematical Thinking: Behavioral and Brain Imaging Evidence. *Science,* 284, 970-974.

Dennett, D. (2000). Making Tools for Thinking. In D. Sperber (ed.), *Metarepresentation: A Multidisciplinary Perspective* (pp. 17-29). Oxford University Press.

Dipert, R. (1993). *Artifacts, Art Works, and Agency*. Temple University Press.

Donald, M. (1991). *The Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Harvard University Press.

Donald, M. (2010). The Exographic Revolution: Neuropsychological Sequelae. In L. Malafouris & C. Renfrew (eds.), *The Cognitive Life of Things: Recasting the Boundaries of the Mind* (pp. 71-79). McDonald Institute Monographs.

Dourish, P. (2001). *Where the Action Is: The Foundations of Embodied Interaction*. MIT Press.

Drayson, Z. & Clark, A. (in press). Augmentation, Agency, and the Spreading of the Mental State. *Consciousness and Cognition*.

Dreyfus, H.L. (2001). *On the Internet*. Routledge.

Dunnell, R.C. (1986). Methodological Issues in Americanist Artifact Classification. *Advances in Archaeological Method and Theory,* 9, 149-207.

Dusek, V. (2006). *Philosophy of Technology: An Introduction*. Wiley.

Eastman, C. (2001). New Directions in Design Cognition: Studies on Representation and Recall. In C. M. Eastman, W. M. McCracken, & W. C. Newstetter (eds.), *Design Knowing and Learning: Cognition in Design Education* (pp. 79-103). Elsevier.

Evans, J.B.T. (2007). Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology*, 59, 255-278.

Evan, J.B.T. & Stanovich, K.E. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, 8(3), 223-241.

Floridi, L. (1999). *Computing and Philosophy: An Introduction*. Routledge.

Floridi, L. (2005). The Ontological Interpretation of Informational Privacy. *Ethics and Information Technology,* 7, 185–200.

Floridi, L. (2006). Four Challenges for a Theory of Informational Privacy. *Ethics and Information Technology,* 8, 109-119.

Floridi, L. & Sanders, J.W. (2004). On the Morality of Artificial Agents. *Minds & Machines*, 14(3), 349-379.

Foer, J. (2011). *Moonwalking with Einstein: The Art and Science of Remembering Everything.* The Pinguïn Press.

Franssen, M., Lokhorst, G.J. & van de Poel, I. (2009). Philosophy of Technology. In *Stanford Encyclopedia of Philosophy* (Fall 2012 Edition), E.N. Zalta (ed.), http://plato.stanford.edu/entries/technology/

Franssen, M., Kroes, P., Reydon, T.A.C. & Vermaas, P.E. (eds.) (2014). *Artifacts Kinds: Ontology and the Human-Made World*. Springer.

Friggs, R. & Hartmann, S. (2012). Models in Science. In *The Stanford Encyclopedia of Philosophy* (Fall 2012 Edition), E.N. Zalta (ed.), http://plato.stanford.edu/archives/fall2012/entries/models-science/

Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford University Press.

Gallagher, S. & Crisafi, A. (2009). Mental Institutions. *Topoi,* 28, 45–51.

Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin.

Giere, R. (2002a). Models as Part of Distributed Cognitive Systems. In L. Magnani, N. Nersessian & P. Thagard (eds.), *Model-based Reasoning in Scientific Discovery*. Kluwer Academic/Plenum Publishers.

Giere, R. (2002b). Scientific Cognition as Distributed Cognition. In P. Carruthers, M. Stich & P. Siegal (ed.), *The Cognitive Basis of Science* (pp. 285-299). Cambridge University Press.

Giere, R. (2002c). Distributed Cognition in Epistemic Cultures. *Philosophy of Science,* 69(4), 637-644.

Giere, R. (2006). *Scientific Perspectivism*. Chicago University Press.

Gooding, D., Pinch, T.J. & Schaffer, S. (1989). *The Uses of Experiment: Studies in the Natural Sciences*. Cambridge University Press.

Griffiths, P. (1993). Functional Analysis and Proper Functions. *British Journal for the Philosophy of Science*, 44, 409-422.

Hacking, I. (1983). *Representing and Intervening*. Cambridge University Press.

Harmon, E. & Nersessian, N.J. (2008). Cognitive Partnership on the Benchtop: Designing to support Scientific Researchers. In *Proceedings of the 7th ACM Conference on Designing Interactive Systems* (pp. 119-128).

Harré, R. (2003). The Materiality of Instruments in a Metaphysics for Experiments. In H. Radder (ed.), *The Philosophy of Scientific Experimentation* (pp. 19-38). University of Pittsburgh Press.

Harris, C.B., Keil, P.G., Sutton, J., Barnier, A.J. & McIlwain, D.J.F. (2011). We Remember, We Forget: Collaborative Remembering in Older Couples. *Discourse Processes,* 48(4), 267-303.

Haugeland, J. (1991). Representational Genera. In W. Ramsey, S. Stich & D. Rumelhart (eds.), *Philosophy and Connectionist Theory* (pp. 61-89). Lawrence Erlbaum Associates.

Heersmink, R. (2011). Epistemological and Phenomenological Issues in the Use of Brain-Computer Interfaces. In C. Ess & R. Hagengruber (eds.), *Proceedings of the International Association for Computing and Philosophy* (pp. 98-102).

Heersmink, R. (2012a). Defending Extension Theory: A Response to Kiran and Verbeek. *Philosophy and Technology,* 25(1), 121-128.

Heersmink, R. (2012b). Mind and Artifact: A Multidimensional Matrix for Exploring Cognition-Artifact Relations. In J.M. Bishop & Y.J. Erden (eds.), *Proceedings of the 5th AISB Symposium on Computing and Philosophy* (pp. 54-61).

Heersmink, R. (2013a). Embodied Tools, Cognitive Tools, and Brain-Computer Interfaces. *Neuroethics,* 6(1), 207-219.

Heersmink, R. (2013b). A Taxonomy of Cognitive Artifacts: Function, Information, and Categories. *Review of Philosophy and Psychology*, 4(3), 465-481.

Heersmink, R. (forthcoming-a). The Metaphysics of Cognitive Artifacts. *Philosophical Explorations*.

Heersmink, R. (forthcoming-b). Dimensions of Integration in Embedded and Extended Cognitve Systems. *Phenomenology and the Cognitive Sciences*.

Heidegger, M. (1962). *Being and Time* (J. Macquarie & E. Robinson, Trans.). SCM Press. (Original work published 1927).

Hilpinen, R. (1992). On Artifacts and Works of Art. *Theoria*, 7, 58-82.

Hilpinen, R. (1993). Authors and Artifacts. *Proceedings of the Aristotelian Society,* 93, 155-178.

Hilpinen, R. (2011). Artifact. In *Stanford Encyclopedia of Philosophy* (Fall 2012 Edition), E.N. Zalta (ed.), http://plato.stanford.edu/entries/artifact/.

Hollan, J., Hutchins, E. & Kirsh, D. (2000). Distributed Cognition: Toward a New Foundation for Human-Computer Interaction Research. *Transactions on Computer-Human Interaction,* 7(2), 174-196.

Holm, S. (2013). Organism and Artifact: Proper Functions in Paley Organisms. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44(4), 706-713.

Holmes, F.L., Renn, J. & Rheinberger, H.J. (eds.) (2003). *Reworking the Bench: Research Notebooks in the History of Science*. Springer.

Houkes, W.N., Kroes, P., Meijers, A. & Vermaas, P.E. (2011). Dual-Nature and Collectivist Frameworks for Technical Artefacts: A Constructive Comparison. *Studies in History and Philosophy of Science Part A*, 42(1), 198-205.

Houkes, W.N. & Vermaas, P.E. (2004). Actions versus Functions: A Plea for an Alternative Metaphysics of Artifacts. *The Monist*, 87(1), 52-71.

Houkes, W.N. & Vermaas, P.E. (2009). Artefacts in Analytic Metaphysics: Introduction. *Techné*, 13(2), 74-81.

Houkes, W.N. & Vermaas, P.E. (2010). *Technical Functions: On the Use and Design of Artefacts*. Springer.

Houkes, W.N. & Vermaas, P.E. (2013). Pluralism on Artefact Categories: A Philosophical Defence. *Review of Philosophy and Psychology*, 4(3), 543-557.

Hughes, R. I. G. (1997). Models and Representation. *Philosophy of Science,* 64, 325-336.

Humphrey, P. (2004). *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford University Press.

Humphrey, P. (2009). The Philosophical Novelty of Computer Simulation Methods. *Synthese*, 169, 615–626.

Hutchins, E. (1983). Understanding Micronesian Navigation. In D. Gentner & A. Stevens (eds.), *Mental Models* (pp. 191-225). Lawrence Erlbaum.

Hutchins, E. (1995a). *Cognition in the Wild*. MIT Press.

Hutchins, E. (1995b). How a Cockpit Remembers its Speed. *Cognitive Science,* 19(3), 265-288.

Hutchins, E. (1999). Cognitive Artifacts. In R.A. Wilson & F.C. Keil (eds.), *The MIT Encyclopedia of the Cognitive Sciences* (pp. 126-128). MIT Press.

Hutchins, E. (2000). Distributed Cognition. In N.J. Smelser & P.B. Baltes (eds.), *The International Encyclopedia of the Social and Behavioral Sciences* (pp. 2068-2072). Elsevier Science.

Hutchins, E. (2005). Material Anchors for Conceptual Blends. *Journal of Pragmatics,* 37, 1555-1577.

Hutchins, E. (2010). Cognitive Ecology. *Topics in Cognitive Science*, 2(4), 705-715.

Hutchins, E. (2013). The Cultural Ecosystem of Human Cognition. *Philosophical Psychology*. DOI: 10.1080/09515089.2013.830548.

Hutchins, E. & Klausen, T. (1996). Distributed Cognition in an Airline Cockpit. In Y. Engstrom & D. Middleton (eds.), *Cognition and Communication at Work* (pp. 15-34). Cambridge University Press.

Ihde, D. (1990). *Technology and the Lifeworld: From Garden to Earth*. Indiana University Press.

Ihde, D. (1991). *Instrumental Realism: The Interface between Philosophy of Science and Philosophy of Technology*. Indiana University Press.

Johnson-Frey, S. (2004). The Neural Bases of Complex Tool use in Humans. *Trends in Cognitive Sciences,* 8(2), 71-78.

Kanare, H.M. (1985). *Writing the Laboratory Notebook*. American Chemical Society.

Kiran, A.H. & Verbeek, P.P. (2010). Trusting Our Selves to Technology. *Knowledge, Technology, and Policy*, 23, 409-427.

Kirchhoff, M.D. (2011). Extended Cognition and Fixed Properties: Steps to a Third-Wave Version of Extended Cognition. *Phenomenology and the Cognitive Sciences,* 11(2), 287-308.

Kirsh, D. (1995). The Intelligent Use of Space. *Artificial Intelligence,* 72, 1-52.

Kirsh, D. (2006). Distributed Cognition: A Methodological Note. *Pragmatics & Cognition,* 14(2), 249-262.

Kirsh, D. (2009a). Problem-Solving and Situated Cognition. In P. Robbins & M. Aydede (eds.), *The Cambridge Handbook of Situated Cognition* (pp. 264-306). Cambridge University Press.

Kirsh, D. (2009b). Projection, Problem Space and Anchoring. In N. A. Taatgen & H. van Rijn (eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2310-2315).

Kirsh, D. (2010). Thinking with External Representations. *AI & Society,* 25, 441-454.

Kirsh, D. (2013). Embodied Cognition and the Magical Future of Interaction Design. *ACM Transactions on Human-Computer Interaction,* (20)1, 1-30.

Kirsh, D. & Maglio, P. (1994). On Distinguishing Epistemic from Pragmatic Action. *Cognitive Science,* 18, 513-549.

Kitching, I.J., Forey, P.L., Humphries, C.J. & Williams, D.M. (1998). *Cladistics: The Theory and Practice of Parsimony Analysis*. Oxford University Press.

Klein, U. (2003). *Experiments, Models, Paper Tools: Cultures of Organic Chemistry in the Nineteenth Century*. Stanford University Press.

Knappett, C. (2010). Communities of Things and Objects: A Spatial Perspective. In L. Malafouris & C. Renfrew (eds.), *The Cognitive Life of Things: Recasting the Boundaries of the Mind* (pp. 81-89). McDonald Institute Monographs.

Knorr Cetina, K. (1999). *Epistemic Cultures: How the Sciences make Knowledge*. Harvard University Press.

Knuuttila, T. & Voutilainen, A. (2003). A Parser as an Epistemic Artifact: A Material View on Models. *Philosophy of Science*, 70, 1484-1495.

Knuuttila, T. (2005). Models, Representation, and Mediation. *Philosophy of Science,* 72(5), 1260–1271.

Knuuttila, T. (2011). Modelling and Representing: An Artefactual Approach to Model-Based Representation. *Studies in the History and Philosophy of Science*, 42, 262-271.

Knuuttila, T. & Boon, M. (2011). How do Models give us Knowledge? The Case of Carnot's Ideal Heat Engine. *European Journal of Philosophy of Science*, 1, 309-334.

Kornblith, H. (1980). Referring to Artifacts. *The Philosophical Review,* 89(1), 109-114.

Kroes, P.A. (2006). Coherence of Structural and Functional Descriptions of Technical Artefacts. *Studies in History and Philosophy of Science Part A*, 37(1), 137-151.

Kroes, P. (2012). *Technical Artefacts: Creations of Mind and Matter. A Philosophy of Engineering Design*. Springer.

Kroes, P. & Meijers, A. (eds.). (2001). *The Empirical turn in the Philosophy of Technology*. Elsevier Science.

Kroes, P. & Meijers, A. (2002). The Dual Nature of Technical Artifacts: Presentation of a New Research Program. *Techné,* 6(2), 4-8.

Kroes, P. & Meijers, A. (2006). The Dual Nature of Technical Artefacts. *Studies in the History and Philosophy of Science,* 37, 1-4.

Kroes, P. & Verbeek. P.P. (eds.) (2014). *The Moral Status of Technical Artifacts*. Springer.

Kurz-Milcke, E., Nersessian, N.J. & Newstetter, W.C. (2004). What Has History to Do with Cognition? Interactive Methods for Studying Research Laboratories. *Journal of Cognition and Culture,* 4(3-4), 663-700.

Laland, K. (2002). Gene-Culture Coevolution. In *Encyclopedia of Cognitive Science* (pp. 268-274). John Wiley and Sons.

Latour, B. (1986). Visualisation and Cognition: Thinking with Eyes and Hands. *Knowledge and Society,* 6, 1-40.

Latour, B. (1992). Where are the Missing Masses? A Sociology of a Few Mundane Artifacts. In W.E. Bijker and J. Law (eds.), *Shaping Technology/Building Society: Studies in Sociotechnical Change* (pp. 225–258). MIT Press.

Lawson, C. (2008). An Ontology of Technology: Artefacts, Relations and Functions. *Techné*, 12(1), 48-64.

Levin, J. (2009). Functionalism. In *Stanford Encyclopedia of Philosophy* (Fall 2012 Edition), E.N. Zalta (ed.), http://plato.stanford.edu/entries/functionalism/.

Levy, N. (2007). Rethinking Neuroethics in Light of the Extended Mind Thesis. *American Journal of Bioethics*, 7(6), 3-11.

Loader, P. (2012). The Epistemic/Pragmatic Dichotomy. *Philosophical Explorations,* 15(2), 219-232.

Magnani, L., Nersessian, N.J. & Thagard, P. (eds.) (1999). *Model-Based Reasoning in Scientific Discovery*. Springer.

Magnani, L. & Nersessian, N. (eds.) (2002). *Model-Based Reasoning: Science, Technology, Values*. Kluwer.

Magnus, P.D. (2007). Distributed Cognition and the Task of Science. *Social Studies of Science,* 37(2), 297-310.

Magnus, P.D. (2009). On trusting Wikipedia. *Episteme*, 6(1), 74-90.

Maguire, E.A., Gadian, D.G., Johnsrude, I.S., Good, C.D., Ashburner, J., Frackowiak, R.S.J. & Frith, C.D. (2000). Navigation-Related Structural Change in the Hippocampi of Taxi Drivers. *Proceedings of the National Academy of Sciences,* 97(8), 4398-4403.

Malafouris, L. (2004). The Cognitive Basis of Material Engagement: Where Brain, Body and Culture Conflate. In E. DeMarrais, C. Gosden and C. Renfrew (eds.), *Rethinking*

Materiality: The Engagement of Mind with the Material World (pp. 53–62). McDonald Institute for Archaeological Research.

Maravita, A. & Iriki, A. (2004). Tools for the Body (Schema). *Trends in Cognitive Science,* 8(2), 79-86.

Margolis, E. & Laurence, S. (eds.) (2007). *Creations of the Mind: Theories of Artifacts and Their Representation*. Oxford University Press.

Meijers, A. (2001). The Relational Ontology of Technical Artifacts. In P. Kroes & A. Meijers (eds.), *The Empirical Turn in the Philosophy of Technology* (pp 81-96). Elsevier Science.

Menary, R. (2007a). *Cognitive Integration: Mind and Cognition Unbounded*. Palgrave McMillan.

Menary, R. (2007b). Writing as Thinking. *Language Sciences,* 29, 621-631.

Menary, R. (ed.). (2010a). *The Extended Mind.* MIT Press.

Menary, R. (2010b). Cognitive Integration and the Extended Mind. In R. Menary (Ed.), *The Extended Mind* (pp. 227-243). MIT Press.

Menary, R. (2010c). Dimensions of Mind. *Phenomenology and the Cognitive Sciences,* (9)4, 561-578.

Menary, R. (2012). Cognitive Practices and Cognitive Character. *Philosophical Explorations,* 15(2), 147-164.

Merleau-Ponty, M. (2006). *Phenomenology of Perception* (K. Paul, Trans.). Routledge. (Original work published 1945).

Michaelian, K. (2012). Is External Memory, Memory? Biological Memory and Extended Mind. *Consciousness and Cognition,* 21(3), 1154–1165.

Michaelian, K. & Sutton, J. (2013). Distributed Cognition and Memory Research: History and Current Directions. *Review of Philosophy and Psychology,* 4(1), 1-24.

Millikan, R.G. (1984). *Language, Thought, and other Biological Categories*. MIT Press.

Millikan, R.G. (1989). In Defense of Proper Function. *Philosophy of Science*, 56(2), 288-302.

Mitcham, C. (1994). *Thinking through Technology: The Path between Engineering and Philosophy*. University of Chicago Press.

Moffatt, B. & Giere, R. (2003). Distributed Cognition: Where the Cognitive and the Social Merge. *Social Studies of Science,* 33(2), 301-310.

Moor, J. (1985). What is Computer Ethics? *Metaphilosophy,* 16(4), 266-275.

Morrison, M.S. & Morgan, M. (eds.). (1999a). *Models as Mediators: Perspectives on Natural and Social Sciences*. Cambridge University Press.

Morrison, M.S. & Morgan, M. (1999b). Models as Mediating Instruments. In M.S. Morrison & M. Morgan (eds.), *Models as Mediators: Perspectives on Natural and Social Sciences* (pp. 10-37). Cambridge University Press.

Neander, K. (1991). Functions as Selected Effects: The Conceptual Analyst's Defense. *Philosophy of Science,* 58, 168-184.

Nersessian, N. (1999). Model-Based Reasoning in Conceptual Change. In L. Magnani, N.J. Nersessian & P. Thagard (eds.), *Model-Based Reasoning in Scientific Discovery* (pp. 5-22). Springer.

Nersessian, N. (2002). The Cognitive Basis of Model-Based Reasoning in Science. In P. Carruthers, S. Stich & M. Siegal (eds.), *The Cognitive Basis of Science* (pp. 133-153). Cambridge University Press.

Nersessian, N. J. (2005). Interpreting Scientific and Engineering Practices: Integrating the Cognitive, Social, and Cultural Dimensions. In M. Gorman, R. Tweney, D. Gooding & A. Kincannon (eds.), *New Directions in Scientific and Technical Thinking* (pp. 17-56). Erlbaum.

Nersessian, N. J. (2006). The Cognitive-Cultural Systems of the Research Laboratory. *Organization Studies,* 27(1), 125-145.

Nersessian, N. (2008). *Creating Scientific Concepts*. MIT Press.

Nersessian, N. (2009). How do Engineering Scientists Think? Model-Based Simulation in Biomedical Engineering Research Laboratories. *Topics in Cognitive Science*, 1, 730-757.

Nersessian, N.J., Kurz-Milcke, E., Newstetter, W.C. & Davies, J. (2003). Research Laboratories as Evolving Distributed Cognitive systems. In R. Alterman & D. Kirsh (eds.), *Proceedings of the 25th Annual Conference of the Cognitive Science Society* (pp. 857-862).

Norman, D. (1991). Cognitive Artifacts. In J.M. Carroll (ed.), *Designing Interaction: Psychology at the Human-Computer Interface* (pp. 17-38). Cambridge University Press.

Norman, D. (1992). Design Principles for Cognitive Artifacts. *Research in Engineering Design,* 4(1), 43-50.

Norman, D. (1993). *Things that Make Us Smart: Defending Human Attributes in the Age of the Machine*. Basic Books.

Norman, D. (1998). *The Invisible Computer*. MIT Press.

Palermos, S.O. (2014). Loops, Constitution, and Cognitive Extension. *Cognitive Systems Research,* 27, 25-41.

Parker, J.N, Vermeulen, N. & B. Penders. (2010). *Collaboration in the New Life Sciences.* Ashgate.

Parsell, M. (2007). The Cognitive Cost of Extending an Evolutionary Mind into the Environment. *Cognitive Processing,* 7, 3-10.

Peck, S.L. (2004). Simulation as Experiment: A Philosophical Reassessment for Biological Modelling. *Trends in Ecology and Evolution,* 19(10), 530–534.

Peterson, D. (1995). Introduction. In D. Peterson (ed.), *Forms of Representation: An Interdisciplinary Theme for Cognitive Science* (pp. 7-27). Intellect Books.

Perini, L. (2005). Explanation in Two Dimensions: Diagrams and Biological Explanation. *Biology and Philosophy*, 20, 257-269.

Perini, L. (2013). Diagrams in Biology. *The Knowledge Engineering Review*. DOI: 10.1017/S0269888913000246.

Peirce, C.S. (1935a). *The Collected Papers of Charles S. Peirce Vol 2*. Harvard University Press.

Peirce, C.S. (1935b). *The Collected Papers of Charles S. Peirce Vol 3.* Harvard University Press.

Pitt, J. (2000). *Thinking about Technology: Foundations of the Philosophy of Technology*. Seven Bridges Press.

Pitt, J (2011). *Doing Philosophy of Technology: Essays in a Pragmatist Spirit*. Springer.

Preston, B. (1998a). Cognition and Tool Use. *Mind & Language*, 13(4), 513-547.

Preston, B. (1998b). Why is a Wing like a Spoon? A Pluralist Theory of Function. *The Journal of Philosophy,* 95(5), 215-254.

Preston, B. (2009). Philosophical Theories of Artifact Function. In A. Meijers (ed.), *Philosophy of Technology and Engineering Sciences* (pp. 213-234). Elsevier Science.

Preston, B. (2013). *A Philosophy of Material Culture: Action, Function, and Mind*. Routledge.

Preston, B. (2014). Ethnotechnology: A Manifesto. In M. Franssen, P. Kroes, T.A.C. Reydon & P.E. Vermaas (eds.), *Artefacts Kinds: Ontology and the Human-Made World* (pp. 145-163). Springer.

Putnam, H. (1975). *Mind, Language, and Reality*. Cambridge University Press.

Radder, H. (ed.) (2003).*The Philosophy of Scientific Experimentation*. University of Pittsburgh Press.

Read, D.W. (2007). *Artifact Classification: A Conceptual and Methodological Approach*. Left Coast Press.

Record, I. (2010). Scientific Instruments: Knowledge, Practice, and Culture. *Spontaneous Generations: A Journal for the History and Philosophy of Science,* 4(1), 1-7.

Reydon, T.A. (2012). Philosophy of Technology. In *Internet Encyclopedia of Philosophy*. Retrieved May 2012, from http://www.iep.utm.edu/technolo/.

Robbins, P. & Aydede, M. (eds.). (2009a). *The Cambridge Handbook of Situated Cognition*. Cambridge University Press.

Robbins, P. & Aydede, M. (2009b). A Short Primer on Situated Cognition. In P. Robbins & M. Aydede (eds.), *The Cambridge Handbook of Situated Cognition* (pp. 3-12). Cambridge University Press.

Roberts, T. (2011). Taking Responsibility for Cognitive Extension. *Philosophical Psychology*, 25(4), 491-501.

Rockwell, T. (2010). Extended Cognition and Intrinsic Properties. *Philosophical Psychology,* 23(6), 741-757.

Rogers, Y. (2004). New Theoretical Approaches for Human-Computer Interaction. *Annual Review of Information Science and Technology*, 38(1), 84-143.

Rowlands, M. (1999). *The Body in Mind: Understanding Cognitive Processes*. Cambridge University Press.

Rowlands, M. (2009). Extended Cognition and the Mark of the Cognitive. *Philosophical Psychology,* 22(1), 1-19.

Rowlands, M. (2010). *The New Science of the Mind: From Extended Mind to Embodied Phenomenology*. MIT Press.

Rubacha, M., Rattan, A.K. & Hosselet, S.C. (2011). A Review of Electronic Laboratory Notebooks Available in the Market Today. *Journal of Laboratory Automation*, 16(1), 90-98.

Rumelhart, D.E., McClelland, J.L., Smolensky, P., Hinton, G.E. (1986). Schemata and Sequential Thought Processes in PDP Models. In D.E. Rumelhart, J.L. McClelland, and the PDP Research Group (eds.), *Parallel Distributed Processing: Explorations in the*

Microstructure of Cognition, *Vol. II., Psychological and Biological Models* (pp. 7-58). MIT Press.

Rupert, R. (2004). Challenges to the Hypothesis of Extended Cognition. *Journal of Philosophy,* 101(8), 389-428.

Scheele, M. (2006). Function and Use of Technical Artefacts: Social Conditions of Function Ascription. *Studies in History and Philosophy of Science*, 37(1), 23-26.

Sheredos, B., Burnston, D.C. Abrahamsen, A. & Bechtel, W. (in press). Why do Biologists Use so Many Diagrams? *Philosophy of Science.*

Simon, H. (1978). On the Forms of Mental Representation. In C.W. Savage (ed.), *Perception and Cognition: Issues in the Foundations of Psychology* (pp. 3-18). University of Minnesota Press.

Simon, J. (2013). Trust. In D. Pritchard (ed.), *Oxford Bibliographies in Philosophy*. Oxford University Press.

Skagerlund, K., Kirsh, D. & Dahlbäck, N. (2012). Maps in the Head and Maps in the Hand. In N. Miyake, D. Peebles & R.P. Cooper (eds.), *Proceedings of the Annual Conference of the Cognitive Science Society* (pp. 2339-2344).

Smart, P. (2012). The Web-Extended Mind. *Metaphilosophy,* 43(4), 446-463.

Søraker, J.H. (2012). Virtual Worlds and their Challenge to Philosophy: Understanding the Intravirtual and the Extravirtual. *Metaphilosophy,* 43(4), 499-512.

Sperber, D. (2007). Seedless Grapes: Nature and Culture. In E. Margolis & S. Laurance (eds.), *Creations of the Mind: Theories of Artifacts and their Representation* (pp. 124-137). Oxford University Press.

Sterelny, K. (2003). *Thought in a Hostile World: The Evolution of Human Cognition*. Blackwell Publishing.

Sterelny, K. (2004). Externalism, Epistemic Artefacts and the Extended Mind. In R. Schantz (ed.), *The Externalist Challenge. New Studies on Cognition and Intentionality* (pp. 239-254). De Gruyter.

Sterelny, K. (2010). Minds: Extended or Scaffolded? *Phenomenology and the Cognitive Sciences,* 9(4), 465-481.

Sterelny, K. (2012). *The Evolved Apprentice: How Evolution Made Humans Unique*. MIT Press.

Suchman, L. (2007). *Human-Machine Configurations: Plans and Situated Actions*. Cambridge University Press.

Susi, T. (2005). In Search of the Holy Grail: Understanding Artefact Mediation in Social Interaction. In B. Bara, L. Barsalou & M. Bucciarelli (eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 2110-2115).

Sutton, J. (2002). Porous Memory and the Cognitive Life of Things. In D. Tofts, A. Jonson and A. Cavallaro (eds.), *Prefiguring Cyberculture: An Intellectual History* (pp. 130-141). Power Publications and MIT Press.

Sutton, J. (2006). Distributed Cognition: Domains and Dimensions. *Pragmatics and Cognition,* 14(2), 235-247.

Sutton, J. (2008). Material Agency, Skills and History: Distributed Cognition and the Archaeology of Memory. In C. Knappet & L. Malafouris (eds.), *Material Agency* (pp. 37-55). Springer.

Sutton, J. (2009a). The Feel of the World: Exograms, Habits, and the Confusion of Types of Memory. In A. Kania (ed.), *Philosophers on Memento* (pp. 65-86). Routledge.

Sutton, J. (2009b). Remembering. In P. Robbins & M. Aydede (eds.), *The Cambridge Handbook of Situated Cognition* (pp. 217-235). Cambridge University Press.

Sutton, J. (2010). Exograms and Interdisciplinarity: History, the Extended Mind and the Civilizing Process. In R. Menary (ed.), *The Extended Mind* (pp. 189-225). MIT Press.

Sutton, J., Harris, C.B., Keil, P. & Barnier, A.J. (2010). The Psychology of Memory, Extended Cognition, and Socially Distributed Remembering. *Phenomenology and the Cognitive Sciences,* 9, 521-560.

Teter, S. & Hartl, F.U. (2005). Protein Folding In Vivo. In: *eLS*. John Wiley & Sons. Retrieved October 2013, from http://www.els.net.

Theiner, G. (2013). Transactive Memory Systems: A Mechanistic Analysis of Emergent Group Memory. *Review of Philosophy and Psychology,* 4(1), 65-89.

Thomasson, A. (2007). Artifacts and Human Concepts. In E. Margolis S. Laurence (Eds.), *Creations of Mind* (pp. 52-73). Oxford University Press.

Thomasson, A. (2009). Artifacts in Metaphysics. In A. Meijers (ed.), *Philosophy of Technology and Engineering Sciences* (pp. 191-212). Elsevier Science.

Thompson, R.B. & Thompson, B.F. (2012). *An Illustrated Guide to Home Biology Experiments: All Lab, No Lecture*. O'Reilly Media, Inc.

Tollefson, D., Dale, R. & Paxton, A. (2013). Alignment, Transactive Memory, and Collective Cognitive Systems. *Review of Philosophy and Psychology*, 4(1), 49-64.

Toon, A. (2012). *Models as Make-Believe: Imagination, Fiction and Scientific Representation*. Palgrave McMillan.

Toon, A. (2013). Friends at Last? Distributed Cognition and the Cognitive/Social Divide. *Philosophical Psychology*. DOI: 10.1080/09515089.2013.828371.

Tribble, E. (2005). Distributing Cognition in the Globe. *Shakespeare Quarterly,* 56, 135-155.

Tribble, E. (2011). *Cognition in the Globe: Attention and Memory in Shakespeare's Theatre*. Palgrave MacMillan.

Tversky, B., Suwa, M., Agrawala, M., Heiser, J., Stolte, C., Hanrahan, P., Phan, D., Klingner, J., Daniel, M.P., Lee, P. & Haymaker, J. (2003). Sketches for Design and Design of Sketches. In U. Lindemann (ed.), *Human Behavior in Design: Individuals, Teams, Tools* (pp. 79-86). Springer.

Tversky, B. & Suwa, M. (2009). Thinking with Sketches. In A. Markman (ed.), *Tools for Innovation* (pp. 75-84). Oxford University Press.

Tweney, R. (1991). Faraday's Notebooks: The Active Organisation of Creative Science. *Physics Education*, 26, 301-305.

Vaccari, A. (2013). Artifact Dualism, Materiality, and the Hard Problem of Ontology: Some Critical Remarks on the Dual Nature of Technical Artifacts Program. *Philosophy and Technology,* 26(1), 7-29.

Vaesen, K. (2011a). Giere's (In)Appropriation of Distributed Cognition. *Social Epistemology*, 25(4), 379-391.

Vaesen, K. (2011b). The Functional Bias of the Dual-Nature of Technical Artefacts Program. *Studies in History and Philosophy of Science*, 42(1), 190-197.

Vaesen, K. (2012). The Cognitive Bases of Human Tool Use. *Behavioral and Brain Sciences,* 35(4), 203-218.

Van Leeuwen, C., Verstijnen, I.M. & Hekkert, P. (1999). Common Unconscious Dynamics Underlie Uncommon Conscious Effect: A Case Study in the Iterative Nature of Perception

and Creation. In J.S. Jordan (ed.), *Modelling Consciousness across the Disciplines* (pp. 179-218). University Press of America.

Verbeek, P.P. (2005). *What Things Do: Philosophical Reflections on Technology, Agency, and Design*. Penn State Press.

Verbeek, P.P. (2011). *Moralizing Technology: Understanding and Designing the Morality of Things*. University of Chicago Press.

Vermaas, P.E. & Houkes, W.N. (2006). Technical Functions: A Drawbridge between the Intentional and Structural Natures of Technical Artefacts. *Studies in History and Philosophy of Science Part A,* 37(1), 5-18.

Von Eckhardt, B. (1995). *What is Cognitive Science?* MIT Press.

Vorms, M. (2011). Representing with Imaginary Models: Formats Matter. *Studies in the History and Philosophy of Science*, 42, 287-295.

Vorms, M. (2012). Formats of Representation in Scientific Theorising. In P. Humphreys and C. Imbert (eds.), *Models, Simulations, and Representations* (pp. 250-273). Routledge.

Vries, M.J, Hansson, S.O. & Meijers, A.W.M. (eds.) (2013). *Norms in Technology*. Springer.

Walter, S. (2010). Cognitive Extension: The Parity Argument, Functionalism, and the Mark of the Cognitive. *Synthese,* 177, 285-300.

Walter, S. (2013). Situated Cognition: A Field Guide to Some Open Conceptual and Ontological Issues. *Review of Philosophy and Psychology*. DOI: 10.1007/s13164-013-0167.

Ward, D. & Stapleton, M. (2012). Es are Good: Cognition as Enacted, Embodied, Embedded, Affective and Extended. In F. Paglieri (ed.), *Consciousness in Interaction: The Role of the Natural and Social Environment in Shaping Consciousness* (pp. 89-104). John Benjamin Publishing Company.

Wegner, D. M. (1986). Transactive Memory: A Contemporary Analysis of the Group Mind. In B. Mullen & G. R. Goethals (eds.), *Theories of Group Behavior* (pp. 185–205). Springer-Verlag.

Weir, A., Chappell, J. & Kacelnik, A. (2002). Shaping of Hooks in New Caledonian Crows. *Science*, 297, 981.

Wheeler, M. (2010). In Defence of Extended Functionalism. In R. Menary (ed.), *The Extended Mind* (pp. 245-270). MIT Press.

Wheeler, M. (2011). In Search of Clarity about Parity. *Philosophical Studies,* 152, 417-425.

Williamson, K. & Cox, R. (2013). Distributed Cognition in Sports Teams: Explaining Successful and Expert Performance. *Educational Theory and Philosophy*. DOI: 10.1080/00131857.2013.779215.

Wilson, R.A. (1994). Wide Computationalism. *Mind,* 103(411), 351-372.

Wilson, R.A. (2004). *Boundaries of Mind: The Individual in the Fragile Sciences*. Cambridge University Press.

Wilson, R.A. (2005). Collective Memory, Group Minds, and the Extended Mind Thesis. *Cognitive Processing*, 6(4), 227-236.

Wilson, R. & Clark, A. (2009). How to Situate Cognition: Letting Nature Take its Course. In P. Robbins & M. Aydede (eds.), *The Cambridge Handbook of Situated Cognition* (pp. 55-77). Cambridge University Press.

Winsberg, E. (2013). Computer Simulations in Science. In *Stanford Encyclopedia of Philosophy* (Fall 2012 Edition), E.N. Zalta (ed.), http://plato.stanford.edu/entries/simulations-science/

Woelert, P. (2010). Human Cognition, Space, and the Sedimentation of Meaning. *Phenomenology and the Cognitive Sciences,* 10, 113-137.

Wright, L. (1974). Functions. *Philosophical Review,* 82, 139-168.

Wylie, A. (2002). *Thinking from Things: Essays in the Philosophy of Archaeology*. University of California Press.

Yeo, R. (2008). Notebooks as Memory Aids: Precepts and Practices in Early Modern England. *Memory Studies*, 1(1), 115-136.

Zhang, J. (1997). The Nature of External Representations in Problem Solving. *Cognitive Science,* 21(2), 179-217.

Zhang, J. & Norman, D. (1995). A Representational Analysis of Numeration Systems. *Cognition,* 57, 271-295.

Zhang, J. & Patel, V.L. (2006). Distributed Cognition, Representation, and Affordance. *Pragmatics & Cognition,* 14(2), 333-341.