Dissertations in Language and Cognition 6

Tim Seuchter

Action-Related Representations

An Action-Based Approach to Grounded Cognition

d|u|p düsseldorf university press

Action-Related Representations

An Action-Based Approach to Grounded Cognition

Tim Seuchter

d|u|p

Hana Filip, Peter Indefrey, Laura Kallmeyer, Sebastian Löbner,
Gerhard Schurz & Robert D. Van Valin, Jr. (eds.)

Dissertations in Language and Cognition

6

Tim Seuchter

2020

Action-Related Representations

An Action-Based Approach to Grounded Cognition

d|u|p

The electronic edition of this publication was made available in open access in January 2022.

ISBN 978-3-11-066378-5
e-ISBN (pdf) 978-3-11-066379-2

Library of Congress Control Number: 2020939663

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

D 61

d|u|p Düsseldorf University Press is an imprint of Walter de Gruyter GmbH Cover illustration: Doris Gerland, Christian Horn, Albert Ortmann
Printing and binding: CPI books GmbH, Leck

dup.degruyter.com

Contents

List of Figures

1 Introduction

1.1 Embodied Cognition, Grounded Cognition and Action-Related Representations

Theories of ‘embodied’ or ‘grounded’ cognition enjoy high popularity in philosophy, psychology and the field of cognitive sciences in general. For at least the last two decades and even more so in the last years, a general trend is noticeable to approach many issues in cognitive science from an ‘embodied perspective’. There is no universally accepted definition for what exactly is referred to by the concept of grounding cognition, and what embodiment comprises exactly; however, some common ground can be identified. Thus, one of the central claims in accounts of both embodied and grounded cognition is that cognition in general depends on the physical constitution of the cognitive system. Cognitive operations, such as thinking, problem solving, memorizing, planning and goal-directed action, can therefore only be completely understood by sufficiently paying tribute to the role of the subject’s body. To be more specific, the claims involve that at least some cognitive processes are based on, or are constituted by processes subserving perception and motor control. Cognition cannot be understood simply as central processing or abstract inference, which is disconnected from action and perception in that these processes are providing merely the input or output faculties to the cognitive system (cf. Wilson 2002). To yield a better understanding of cognitive processes, a more fundamental role has to be assigned to perception and action. This can refer to different aspects: perception can be analyzed as an active process which crucially involves movement and the body of the subject (Merleau-Ponty 1945/2012[1]; O'Reagan & Noë 2001). Cognition can be described as being grounded in concrete perceptual symbols that are stored as perceptual representations becoming reactivated at later occasions (Barsalou 1999). Understanding language can be explained with simulated or reenacted motor knowledge (Lakoff & Johnson 1980, 1999; Pulvermüller 2005). Embodied cognition can also be understood as using one's body for problem solving, such as finger counting for solving mathematical problems (Fischer & Brugger 2011).

Accounts of embodied and grounded cognition are in opposition to views subsumed under the label 'computationalism'. A common claim of computational accounts is that cognition can be best described in analogy to a digital computer: perception generates input, a central computing unit processes the input information and generates an output in terms of a motor command. These three domains are strictly separated from a computational point of view and have all their own underlying codes and algorithms. Prominent advocates of this conception of cognition have been, among others, Fodor (1975, 1983), Newell and Simon (1976), Pylyshyn (1984) and, more recently, Edelman (2008), their views have had a major impact on the way scientific research conceived of cognitive processes. Computational views of cognition are the logical outcome of the endeavor of modelling truly intelligent artificial systems. Treating cognition as complex manipulations of physical symbols implies that, in principle, cognitive processes can be implemented in computers and machines of given sophistication, as the general processes are hardware independent and simply require enough computational power. The rise and success of early research of artificial intelligence has thus been among the reasons why this view of cognition became notorious. Computationalism most often embraces 'strong representationalism', the view that cognitive processes are syntactical or algorithmic manipulations over central units which are

1 Throughout this work, I will refer to the 2012 translation of the 'Phenomenology of Perception', which has been first published in 1945. The latest translation is a substantial revision and improvement to former translations, which is why I will only refer to this edition. To avoid confusion and providing misleading historical contextualization, the reference will always include the year the first edition was published.

described as (mental) representations. This view is, to some degree, analogous to the software/hardware distinction in digital computers, also assuming that mental representations are discrete states that can be syntactically combined and generate meaningful semantic content this way.

The central criticism of accounts of embodied and grounded cognition is that the role of perception and action for cognition has been entirely misunderstood and, due to this misconception, seriously neglected. Thus any theory of cognition has to account for the roles perceptual input and motoric output play in the cognitive system, as these domains largely overlap and cannot be treated separately. As mentioned above, theories of embodied cognition and grounded cognition can differ significantly in their central premises and their explanatory scope (for detailed descriptions, see Anderson 2003; Wilson 2002).[2] The term 'embodiment' is generally used in a much broader way than 'grounded cognition'. Embodiment generally assigns an important role to the body in understanding, explaining and analyzing cognition in general (cf. e.g., Merleau-Ponty 1945/2012; Gallagher 2005; also, to some extent, Lakoff & Johnson 1999) often without being explicit on what the body actually is and which bodily processes it includes or excludes. Accounts of grounded cognition often have a more specific focus on the exact role sensory and motoric representations play for cognitive processes (cf. e.g., Barsalou 1999, 2008; Glenberg & Kaschak 2002; Jacob & Jeannerod 2003; Zwaan 1999). Thus, the notion of grounded cognition is already implying a commitment to representationalism, which entails that cognitive processes are crucially involving (mental) representations, such as perceptual, motoric, conceptual and abstract representations. Embodiment, on the other hand, can have a broader reading, and although representationalist accounts of embodied

2 This distinction offered here between embodied and grounded cognition is by no means exhaustive, but only points to different foci in the different approaches. For dividing the fields adequately, exact definitions would have to be introduced first, which would take a lot more time than can be spent here. Besides, many other labels for similar and different approaches exist, such as enactivism, embedded cognition, situated cognition etc., which will also not be differentiated further. For the current purpose, theories that ground representational content in sensory and motoric representations are subsumed under the label of 'grounded cognition', whereas 'embodiment' is interpreted as a broad notion that considers a crucial role of the body for cognition while not necessarily being committed to representationalism.

cognition exist, many of them are modelled in terms of anti-representationalist frameworks, such as dynamic systems theory (cf. Chemero 2009; Hutto & Myin 2012; Smith 2005).

The notion of cognition that does not consist of representations as building blocks is problematic for many reasons and a thorough discussion of all the difficulties that, e.g., dynamic systems accounts have, would go beyond the scope of this project, which is why I will only briefly mention some of the main problems in the next section.

For now, the central premise of this book will be that cognition crucially involves representations. Representations are taken to be contentful states of the cognitive system and thus exemplify intentionality. They can be about the world, about other representational states, or contain information about the subject's body and thus be the vehicle for low-level cognitive processes, which might not even be representations on their own. Furthermore, representations can be modal-specific (such as a purely auditory representation) or can be multi-modal and contain information from, e.g., different sense modalities. Representations are structured entities and can vary significantly in complexity. Among the most complex structured representations are conceptual representations, which feature in higher-level cognition such as thinking and linguistic abilities. Most known cognitive functions are supposed to rely on processing representations, which in turn are crucially involved in perception, memory, goal-directed action, imagination and logical reasoning, among others.

Most theories of grounded cognition will accept these premises, moreover, they will make an attempt to solve the so-called 'symbol grounding problem' (cf. Harnad 1990; Searle 1980), or a version thereof (cf. Barsalou 2008). According to advocates of the grounding problem, standard computational cognitive theories, which take cognition to be manipulation of amodal representations whose content can be defined entirely in terms of their syntactic features and functional role in cognitive operations, face the problem of accounting for representational content without becoming circular or having to introduce controversial innate concepts or modules (e.g. Fodor 1983). As cognition, thus described, is mere manipulation of "meaningless symbols", these representations can only refer to other syntactically defined representations and it becomes mysterious at which stage actual content, as exemplified by a cognitive state, which represents aspects of the external environment, should arise. Accordingly, representational content has to be grounded in something, and theories of grounded cognition argue that representations are grounded in perceptual experience and motor action-output of the subject. As perception and action are taken to be based on representations itself, the content of cognitive representations is understood to be derived from sensory and motoric representations, whose content in turn is a result of fundamental structures underlying visual and motoric processing. Perceiving and interacting with the environment establish the original content-generating relations that cognition is taken to be grounded in, thus no circularity worries arise and the meaning of mental representations can be accounted for without problematic assumptions about the genesis of representational content.[3]

Many theories of grounded cognition have a strong focus on perception, i.e., grounding representational content of concepts in perceptual representations (cf. Barsalou 1999; Prinz 2005). These approaches account for object concepts, such as 'chair', on the basis of former perceptual encounters with chairs, which are stored in memory and form (with the assistance of cognitive abstraction mechanisms of some description) a representation of the category 'chair'. Leaving out the details and individual differences of the different accounts, the basic idea is that perceptual encounters are the source for representational content and deploying these representations in later cognitive processes is best understood as some form of simulation or reenactment of the original encounter. Accounts of grounded cognition that focus more on the role of action for representational content and cognitive abilities are often less specific about the nature of the relevant representations involved and how they were generated. Generally, most accounts of grounded cognition hold that the main

3 Theories of grounded cognition have other explanatory merits besides accounting for the meaning of representations in terms of sensorimotor experience, such as providing an explanation for conceptual flexibility and conceptual development. By taking different experiential histories of individuals into account, the differences in meaning of concepts and the changes in meaning can be explained much more easily than amodal computationalism accounts could be (Pecher et al. 2010).

function of cognition is to guide and enable interaction with the environment, and therefore take action to be central to representation. (cf. Pecher et al. 2010).

Glenberg & Kaschak (2002) provide one of the most prominent and influential account of grounding linguistic cognition in action, provide convincing evidence for the 'action-sentence compatibility effect', which describes sentence understanding as based on bodily action. For example, the movements involved in reaching out and giving an object to another subject are part of the understanding of the linguistic expression 'He gave her the pizza' (cf. Glenberg & Kaschak 2002). Linguistic comprehension is taken to be grounded in bodily action, to involve action or to be based on action. Aside from referring to Piaget's (1954) idea that the concept of causality is developed on the basis of the child's registration of causal impact on her environment and O'Reagan & Noë's (2001) idea of acquiring sensorimotor contingencies as the basis for perceptual knowledge, Glenberg & Kaschak are unspecific on the nature of the 'action-grounding'. Although the general focus of grounding is on action, the notion of action is not explicated and thus it is unclear if action in their account refers to the subject's actual action-skills, descriptive knowledge of actions, the subject's ability to imagine actions or simulating formerly performed actions mentally. Other accounts, such as, Borghi (2004) provide more detailed interpretations of what grounding in action could possibly mean. Borghi describes the action aspect of the grounding relation either as stored patterns of motor-cortical activation on the original encounters with object interaction (cf. Borghi 2004, 70f), or as encodings of possible action patterns regarding the subject's environment. Pulvermüller (2005) describes brain mechanisms that are involved and correlated to processing action verbs, such as 'kick', 'pick' and 'lick'. These findings show a significant overlap in cortical regions activated in actual action execution and processing linguistic information expressing the very actions. According to Pulvermüller's results, grounding the meaning of linguistic expressions of actions can be interpreted as actually reenacting these actions on the neuronal level. Similar findings are reported from Chao & Martin (2000), which identify neuronal activation during tool use with similar activation in cognitive tasks involving the same tools, such as viewing and naming tools.

Although these and other empirical studies provide evidence for the role of the motor-cortical areas in other cognitive tasks and thus strengthen the plausibility of grounded cognition theories, a clear understanding of the role and nature of action for grounding cannot yet be established on the basis of neurobiological findings alone. Action cannot be reduced easily to motor-cortical activation patterns, as action in general involves a variety of different aspects, of which motor activation is only one of many. Thus, action is generally held to be goal-directed behavior, which implies that action representation involves goal states, such as objects being action-targets. Furthermore, actions are highly contextual, which means that most actions are situated in an environment, which might be different at each instance. This implies that actions cannot be represented without crucially taking into account the contextual features, such that there is little sense for representing isolated actions without situations in which they are meaningful. A kicking action without an object to kick is a rather abstract movement pattern and unlikely to be the prototypical action representation for 'kicking'. This leads to a more specific notion of action as relatum for grounding cognition: kicking a ball is an action a subject might have performed at a given occasion, which enables this subject to retrieve the stored memories of this very kicking action. The neuronal representation of 'kicking' thus refers to a specific kicking action in each individual, until this subject has formed an abstract concept of 'kick'. Of course, in all instances of kicking representations, quite likely neuronal motor activation will occur, however, for a theory of grounding, it has to be specified what exactly this neuronal activation pattern stands for: it could be an individual's personal experience of kicking a ball, be simply based on observation of another subject kicking a ball, or it could represent an abstracted action category of 'to kick' that is correlated to the neuronal pattern. The evidence, presented by Pulvermüller and others, clearly suggests a functional involvement of some motor cortical areas in language processing and is thus a strong support for the idea of grounded cognition, though yet, it cannot provide the full meaning of the idea of 'grounding cognition in action', as the notion of action in all its relevant dimensions cannot be specified by motor cortical activation patters alone.

This brief overview shows that more thorough analysis is needed for developing a better understanding of the action-aspect in the grounding relation. The notion of action, as well as the idea of sensorimotor representations, can be understood in many different ways as they involve a variety of different aspects, which needs to be systematized to develop a theoretically applicable concept. Interestingly, this appears to be different for perceptual representation, having been the object of enquiry in various scientific disciplines for centuries. Despite all controversies about the nature of perception, it seems that it is much easier to agree on a viable notion of perceptual grounds for cognitive abilities than it is for the notion of action. While most accounts of grounded cognition agree that action is important to representation, they rely at the same time on a rather superficial analysis of what the action part of representation is.

Possible candidates for grounding cognition could thus be neuronal activation of motor cortical areas in action planning and execution, which becomes simulated or re-enacted at other occasions. Merely mentally imagining actions, in terms of movement successions, could be relevant for conceptual knowledge, and thus some cognitive abilities could be grounded in motor imagery. Cognition could also be grounded in either concrete or merely potential movements, such that thinking about action involves representations of the subject's planned or executed movements. Finally, there is general need for clarification of what the concept of action referred to implies. A common way to distinguish actions is by describing simple actions, such as reaching for a glass, in contrast to complex actions, such as drinking, which can involve many simple actions, such as reaching for a glass after filling it with water etc. An even more complex action is giving a toast at a reception, which, among others, also involves the simple action of reaching for a glass. Hence, if action is identified as possible grounds for cognition, it has to be made clear which aspects of action can be relevant for grounding.

Attempting to resolve some of these issues, a central claim of this thesis will be that action can only be grounds for cognition by introducing a kind of representation that captures the relevant action aspects. Action-related representation is a kind of representation that represents features of the environment in terms of possible movements. Possible movements in turn are represented by motor-cortical activity, involved in action planning and action execution. Simple action-related representations can be distinguished from complex action-related representations. Simple actions involve simple movements, which correspond to certain bodily features, such as in reaching, pointing or grasping movements. Thus, a simple action-related representation represents features of an object, such as its size, width, or distance, in terms of simple possible movements, such as the grasping movement one has to execute to pick up the object. This structural simplicity of action-related representations is the main reason for their ability to function as grounds for cognition. More complex action-related representations can be described as being built upon simple action-related representations and being of a more complex structure, as well as representing features of the environment in a more systematic way. An example for a more complex action-related representation would be representing an object, such as a bottle, in terms of the different action-possibilities it allows for. A bottle can be used for all kinds of different purposes; it can be a container for liquids, a door stopper, a hammer or a paper weight, depending on the situational requirements. This kind of action-related representation involves a lot more practical (or theoretical) knowledge of the subject, which captures the idea that with the changing set of behavioral skills of subjects, so too does their capability to represent something as possible action. Action related representation is thus able to describe development from very basic action skills to more complex set of skills that are related to features of the environment.

The idea that there is an "action-mode" of representing one's environment can be found in various discourses throughout the last 100 years. Very prominently, Merleau-Ponty (1945/2012) addressed the intentionality of the body and the idea that the body's existence is defined by being in a practical field, which, similar to a visual field, locates the subject in a space of action-possibilities, always involving the subject's body schema. Gibson's (1986) notion of 'affordances' is similar in spirit, but puts even more emphasis on the idea that fundamentally, all animals perceive their environment in terms of what the environment 'affords', meaning that action possibilities exist only relative to the animal. Although both Merleau-Ponty and Gibson's accounts are decidedly anti-representational, they still are of utmost importance for the present discussion, as they promote the general idea of an action-oriented world approach that is the transcendental condition enabling all higher-order cognitive skills.

More contemporary accounts that focus on the role of action in representation or describe action-possibilities as a mode of representation are the concept of 'pushmi-pullyu representations' (Millikan 1996), 'interactive representation' (Bickhard 1999), 'visuomotor representations' (Jacob & Jeannerod 2003) or 'causal indexicals' (Campbell 1993, 1994), just to name a few. These accounts and the others that will be discussed in the following chapters analyze and focus on different aspects of action, while sharing the general idea that representations that have the function to guide or control actions are among the fundamental representations for cognition.

Action-related representations, as described and specified in the following chapters, are plausible candidates for grounding cognition as they bring together the two central elements of perception and action. By defining this type of sensorimotor representations, which encodes sensory input in a movement format, a basic level is defined from which more sophisticated representations can be derived. Representations that represent in terms of action, i.e. in terms of movements, are composed of very basic units that cannot be analyzed much further – basic sensory and motoric activations are at the core of action-related representation and correspond to basic skills of subjects, which in turn provide the foundation for other skills and cognitive abilities. Moreover, action-related representations can be used to demonstrate how more abstract cognitive abilities can also be grounded in action. Although no fully elaborated theory of cognitive abstraction, covering all aspects of abstract mathematical cognition or abstract concept in humans, can be presented in this thesis, it will be argued that an abstraction mechanism can be identified on the basis of action-related representation. This mechanism allows for representations that only represent highly contextual features to become more general in their signification and can be the basis for classification operations and generalized object representation. The abstraction mechanism thus described is able to illustrate how a subject acquires simple object concepts by interacting with these objects on the basis of basic action-related representations. This developmental aspect is of great value to the debate of grounding cognition, as defining the grounds for cognition also needs a developmental story for how to proceed from these grounds. Analyzing action-related representations can contribute to a better understanding of the development of representational skills in animals and human babies.

Action-related representations, in their simplest form, can be attributed to a great variety of species. From an evolutionary perspective, it thus seems very plausible to depart from basic representation of possible actions in explaining the development of cognition. The idea of grounding cognition in action accounts for the fact that living beings are primarily acting beings – creatures that interact with their environment, in more or less flexible ways and with varying degrees of sophistication. It captures the idea that vision is foremost for motor control, implying that animals perceive in order to guide and control their movements and interactions. Although there might be additional evolutionary explanations for the development of complex sensory systems, it is plausible to assume that the main purpose of the sensory systems is to enable animals to interact successfully with their environment. Organisms that are incapable of self-produced movements, such as sea anemones, hardly need a complex sensual system, as they are not capable of flexible behavior anyway. So all kind of interaction relies on sensorimotor representation and these, according to the grounded cognition framework, build the basis of, and might even be constitutionally involved in, low-level as well as higher-order cognitive processes.

1.2 Representationalism

As mentioned earlier, embodied cognition as a general research paradigm is not restricted to presupposing (mental) representations as the central elements of cognitive processes. In fact, one way of attempting to overcome the problems of standard computational accounts of cognition was to abandon representation generally and focusing on the dynamic aspects of cognitive processes and the coupling relations of cognitive systems and their environment. Thus, Brooks (1991), who holds that intelligence is essentially embodied, also claims that representations are the wrong units in accounting for intelligent behavior, resulting in the infamous slogan ‘‘The world is its own best model” (Brooks 1991, 15). Beer (2003) claims that

> rather than assigning representational content to neuronal states, the mathematical tools for dynamical systems theory are used to characterize the structure of the space of possible behavioral trajectories and the internal and external forces that shape the particular trajectory that unfolds. (Beer 2003, 210)

Chemero (2009), arguing for a new radical embodied perspective on cognition that entirely dispenses with mental representations, holds that mental representations are mere theoretical postulations and should be substituted with a dynamical framework of action, perception and the environment. In particular, his argument against representations is an epistemological claim: explaining cognition does not need the positing of representation and thus representation should be dismissed. The metaphysical claim that there probably are no representations in cognitive systems is an indefensible claim as it does not involve a scientific hypothesis but is rather the product of philosophical speculation (cf. Chemero 2009, 67).

The most important and most widely accepted argument against representations is thus that cognition can be explained without presupposing representations, which makes representations stipulated, theoretical entities without extra explanatory value. This argument comes in many shapes and variations, which cannot be all presented in full detail here (e.g. Beer 2003; Brooks 1991; Chemero 2009; Gibson 1986; Thelen and Smith 1994; van Gelder 1995). I will simply assume that the argument in its general form (from which all other versions are derived) is the strongest case that can be made against representationalism. Against these claims, I will provide some reasons why representations are important for explaining cognitive processes and abilities, which renders them superior to other possible explanations that dismiss representations.

1.2.1 Narrow or Wide Definition of Representation

First of all, it has to be noted that the term 'representation' can have different definitions and, according to a broad understanding can denote all sorts of cognitive states. If the notion of representation is allowed to be sufficiently broad, cognitive states and processes featuring in various anti-representationalist accounts can in fact be interpreted as meeting all the criteria representations as cognitive states should have and thus no real conflict arises – it simply turns out to be a mere difference in labelling the same thing. Thus, when Gibson (1986) speaks of the perceptual system resonating to information specified in the ambient light array, the resonance induced by the environmental stimuli can be interpreted as representing information about the animal's environment. As Gibson does not further specify what exactly he means by 'resonance', interpreting it in representational terms is a valid option.

If representations are defined as necessarily discrete cognitive states, as Van Gelder (1995) does, then every account stressing the dynamicity and analog nature of cognitive states will dismiss representations. However, cognitive representations are by no means restricted to representing only static content or being static in nature generally, neither conceptually nor empirically. Many accounts treat representations as potentially dynamic, temporal representations integrating past, present and future events (Huber 2012), or body-related representations such as the body schema (Gallagher 1995), which can be interpreted as integrating constantly changing information about one's bodily constitution and practical skills. In fact, it seems rather odd to define representations as static and discrete entities, while at the same time most cognitive operations are taken to be inherently dynamic and interactive. Thus, generative models of representation hold that "representational capacity and inherent function of any neuron, neuronal population or cortical area is dynamic and context sensitive" (Friston & Price 2001).

Without going into detail, it seems fair to claim that any viable notion of representation should be able to account for dynamic aspects in cognition as well as the cognitive system-environment interaction, thus dynamicity is not a limitation of representation, but an aspect thereof. The general claim of this section is that a misleadingly restrictive definition of representation is used by many anti-representationalist accounts. However, representation does not have to be interpreted in such a restrictive way, and by accepting a more flexible and broader definition, many of the problems associated with representations disappear, such as Gibson's homunculus criticism (see ch. 3 for more details) or the alleged static nature of representations. A broader definition of representation, e.g., one solely in terms of functional roles can be applied to many cognitive states that are taken to be explanatorily relevant by anti-representationalists.

1.2.2 Higher-Order Cognitive Functions Presuppose Representations

A central aim for any theory of cognition is to explain behavior on the basis of underlying cognitive operations. Behavior varies greatly in complexity and level of sophistication: from simple reflexes such as ducking one's head due to a fast approaching object, to drawing an object in art class to memorizing all American presidents and their periods of governance. These behaviors rely on different cognitive resources and therefore quite plausibly require explanations of different levels of complexity. One way to account for more sophisticated behavior that involves learning, memory and other complex skills relies on representations and representational content that essentially enables and drives these abilities. Thus, this claim is fairly simple and general: it is very implausible to find a convincing explanation for sophisticated behavior merely in terms of dynamically coupled systems. Instead, the cognitive states involved in sophisticated abilities need to be individuated by their content in addition to any dynamic process description they might feature in. Activities, such as catching a ball, might as well be readily describable in terms of dynamic systems, but reducing all possible behavioral complexity to ball-catching-scenarios is highly implausible and thus representations will sooner or later have to enter the picture.

Moreover, representational explanations have the further advantage of being able to account for perceptual illusions and phenomenal appearance in general. How should an explanation of perceptual states that have a content deviating from the actual properties of the perceived situation look like, without relying on representational content? As features of objects, according to the representational view of perception, are not directly perceived but instead represented by the subject, the difference between represented and actual features can be accounted for (see ch. 3 on the problems of direct perception). The same holds for illusionary or imagined features, which can also be explained by top-down effects involving stored representational knowledge. Even if a bottom-up explanation for some perceptual phenomena can plausibly be given, this does not rule out the explanatory advantage of representations for many other cases.

An interesting case of higher-order cognition is mental imagery, the ability to picture state of affairs and processes before one's "inner eye". Mental imagery is most plausibly explained on the basis of reactivated, reenacted stored representational knowledge. As studies have shown, imagining an action underlies the same biomechanical constraints as real, executed action. For instance, subjects were found to be generally unable to imagine faster movements than they could actually produce (Jeannerod 2007). Thus, motor imagery is using the same cognitive resources as actual executed action, which can be best explained if some of the underlying representations used are shared by both processes. If imagery is explained on the basis of actual execution of the same operations, then it is very plausible to assume that the representational basis is shared. Thus, representations are the core elements for low-level and higher-level cognitive abilities, the latter being derived from the former. It follows that mental imagery is not only a cognitive ability that is hard to explain without allowing for representations, but the underlying processes can also be interpreted as being representational.

So far, no convincing account of higher-order cognition has been presented that can account for the full range of cognitive phenomena we want to explain while completely dispensing with representations. Higher-order cognitive abilities are therefore most plausibly involving representations. As representations underlying complex cognition have to be rooted in more basic cognitive processes, a connection between complex and basic-level representations has to be established. One endeavor of this work is to show how representations enabling higher-order cognition systematically develop on the basis of basic-level representations.

1.2.3 A Special Case: Action-Related Representations

The last aspect is that action-related representations have a special status in the cognitive sciences and are widely accepted by both representationalists and anti-representationalists about certain aspects of cognition. Thus, Chemero (2009) states that

> Action-oriented representations differ from representations in earlier computationalist theories of mind in that they represent things in a nonneutral way, as geared to an animal's actions, as affordances. Action-oriented representations are more primitive than other representations in that they can lead to effective behavior without requiring separate representations of the state of the world and the cognitive system's goals. (Chemero 2009, 26)

This special feature of action-related representations has been embraced by embodied cognitive scientists, as it is a less controversial notion of representation while proving substantial explanatory value. Even though Chemero's point is to argue against the assumption of representation in general, the passage supports the aim of the current project. In being primitive, action-related representations can be the grounds for the further development of more complex representations. It would be misguided to assume that only primitive representations can exist and the rest of cognition has to be explained on a different basis. It is representation all the way down, as Fodor infamously stated. By taking action-related representation to be a special case of representation that does not have the primary function of representing neutral facts or state of affairs, but instead guiding behavior, a much more adequate developmental picture arises. The ability of flexibly adjusting behavior is of major advantage for organisms, as they can adapt more readily to changes in their environment. Representation-based behavior control is the most plausible mechanism for flexible behavior, because it allows for explaining why sometime the presence of certain condition lead to executed behavior, while at other occasions the same stimuli do not elicit behavioral response or response of a different kind. By this move, behavior becomes detached from environmental stimuli and more flexible behavioral reactions to environmental situations are possible.

The claim that is made here is that as soon as one allows for primitive, behavior guiding representations, the door is wide open for accepting representations as the basis for cognition in general. As the central line of argumentation in this thesis concerns the nature of simple action-related representations and how they contribute to cognitive abilities of different degrees of complexity, it will become clear that action-related representation is a promising starting point for theories of grounded cognition in general. Accepting action-related representation is easy for many cognitive scientists, and I will show how this assumption coheres with explaining other aspects of cognition on a representational basis as well.

1.3 Overview

The aim of this work is to develop an account of action-related representation that captures the cognitive processes underlying interactions with one's environment while at the same time providing a possible foundation for grounded cognition. To achieve this, other accounts that emphasize the importance of action for various cognitive abilities will be discussed first. To start with, Merleau-Ponty's (1945/2012) notions of 'body schema' and 'motor intentionality' will be presented in chapter 2. Merleau-Ponty has a non-representationalist understanding of cognition, nevertheless his notion captures the central idea that an action-orientation is essential to living beings and neither the body nor perception can be fully understood without taking into account that living beings are foremost interactive beings. In this sense, Merleau-Ponty can be seen as an important precursor to the contemporary debate about embodied cognition by arguing that the body and its capacity for actively engaging in the environment are central to understanding all other cognitive operations.

In chapter 3, one of the most important and possibly most controversial contributions to the psychology of action in the last 50 years, namely Gibson's (1986) concept of affordances will be critically analyzed. In developing an ecological psychology, Gibson sought to overcome the problems he saw in contemporary accounts of cognition of his time. His concept of affordances is central as it is understood as transcending the objective-subjective distinction by promoting a version of direct realism. According to this framework, opportunities for actions (affordances of the environment) are perceived directly by animals, which entails that no mediating cognitive processes or operations contribute to perception in general. As perceiving action possibilities relates both properties of the environment and properties of the animal, Gibson's account is intended to explain how different animals perceive different action opportunities without representing them mentally or otherwise, but instead solely by their physical constitution that is related to environmental features. It will be shown that Gibson's central ideas provide valuable insights to the role the body plays in cognition and especially in determining action possibilities of animals, but that central aspects of his account have to be substantially revised in order to overcome the severe problems arising from Gibson's ontological commitments.

Chapter 4 is about the claim that representations underlying action planning and generation are inherently egocentric and thus implicitly represent the agent. Central to this discussion is Campbell's (1993, 1994) notion of 'causal indexicals', representations with direct reference to the representing subject and immediate consequences for the subject's actions. Analyzing this notion (among other similar accounts) of implicit self-representation will show how self-representing aspects are an essential part of even the most basic action-related representations and are thus fundamental for developing more sophisticated concepts of agency and the self. Causal indexicals furthermore have the potential to establish a basis for preconceptual object representations based on the representing subject's abilities and physical constitution.

An important group of representational theories claims that the main function of representation in general is action-guidance. This claim, elaborated in chapter 5, opposes the view that representing one's external environment has the purpose of providing the subject with "neutral" factual information about the environment. By emphasizing that action-guidance is the primary function of cognitive representations, enabling goal-oriented behavior for all sorts of organisms, these teleo-functional views argue for an evolutionarily adequate approach explaining the origin of representation. Claiming that action-guiding representations are developmentally basic, more complex and sophisticated representations are understood as emerging on the basis of their simple predecessors.

Chapter 6 focusses on the neuro-functional mechanisms enabling the visual processing of action-related features of the environment. Starting from Milner and Goodale's (1995) famous and well established 'dual pathway hypothesis', which identifies two functionally distinct cortical regions, processing either information for object identification or information useful for object interaction, the focus of the remaining chapter will be on the more refined account of Jacob and Jeannerod (2003). The latter confirm the dual pathway hypothesis in general but move on to identify two ways of processing visual information: pragmatic processing involves processing information relevant for action generation, while semantic processing leads to factual and conceptual knowledge about the world. An important aspect of pragmatic processing is the distinction between low-level pragmatic representations that allow for simple interactions with objects in one's environment, and the more complex higher-level representations at work in more complex action scenarios. While the former process is understood to be mostly unconscious and automatic, the more complex pragmatic representations involve structured object information and are more easily consciously accessible. This distinction supports the claim that more abstract representations are developed on the basis of representations enabling simple actions and are thus an important contribution to the idea of grounding cognition in action-related representations.

So far, the accounts presented focused on rather specific aspects of action representation and the role for certain cognitive functions. In chapter 7, two accounts will be presented that explicitly claim that interaction is an essential condition for the development of cognitive abilities in general and the development of intelligent cognition, such as thinking in particular. Piaget (1977) addresses the role of action for the development of thinking, claiming that the subject-object distinction develops on the basis of the child's increasingly systematic interaction with the environment, as well as the development of object concepts and abstract thought processes. Bickhard (1999) building on these Piagetian assumptions, moves on to argue for a general account of interactive representation that explains the development of cognitive abilities for all sorts of living and artificial cognitive systems. One central claim is that interactive representation is the most basic kind of representation while at the same time emerging from non-representational states, thus not presupposing existing representation. This conception of representation genesis is fundamentally grounded (as per general idea of grounded cognition), as Bickhard defines the lowest level of representations in terms of processes that are by themselves not representational and thus of such a simple structure that they exist across all species and artificial systems. The basic elements of interactive representations are motoric output and the respective feedback information, determining further states of the system. Object representations are, according to this view, the outcome of multiple interaction situations, resulting in bundled opportunities for action.

On the basis of interactive representation, a theory of action-related representation can be developed that accounts for structured and stable representations, such as object representations in terms of possible interactions. In chapter 8, the general account of action-related representation will be introduced. It is both a summary of accounts discussed in the previous chapters as well as a synthesis of those aspects that appear central to action representation, action generation and action guidance. The core features of action-related representations are egocentricity, goal-directedness and their being basic in nature. Intentions for actions are sufficient conditions for action-related representation, but not necessary ones, entailing that action-related representations are logically independent form intentions. Action-related representations, in their most prevalent form, are automatically generated and represent features of the environment in terms of simple movements. Simple movements in turn are determined by the physical constitution of a subject. An object is thus represented in terms of a reaching, pointing or grasping movement of a specific subject. This simple way of representing features of the environment relates bodily features of subjects to environmental features, similar to Gibson's (1986) idea of affordances. Action-related representations are different from Gibsonian affordances, as they are highly subjective inner models of external environments, thereby avoiding the difficulties that arose with Gibson's ontological commitments to a direct realism.

In chapter 9, a theory of cognitive abstraction mechanisms is developed on the basis of the general account of action-related representation. If action-related representation is supposed to be a fundamental aspect in cognitive development, the development of more complex and abstract representations on the basis of simple action-related representations has to be accounted for. Accordingly, abstraction is defined as the extension of a subject's frame of reference, from purely egocentric and context bound to allocentric and context independent. Another important aspect of abstraction is the transition from an implicit self-representation of the agent and environmental features to an explicit representation of environmental features and oneself as an agent. Being able to explicitly represent oneself as an agent is a condition for developing a concept of self and self-consciousness. These two aspects of abstraction correlate with the development of action skills, resulting in broader behavioral repertoire and increasingly flexible behavior. From unsystematic interaction with objects, e.g. in a human baby's first weeks, abilities such as object permanence are developed over time in cognitive and motor development, thus enabling the formation of stable object representations. Thereby, a transition from implicitly representing an object's features ('is graspable for me now') to an explicit representation of some of the object's features, are analyzable parts of the representation, derived from former interactions. The ability for perspective taking, as exemplified by the false-belief test (Wimmer & Perner 1983), points to a general ability of thinking of other subjects as agents with intentions and goals. The main claim following from these findings is that by interacting with the environment, subjects first come to develop an explicit representation of themselves as distinct from the world, manifesting in the fundamental subject-object distinction that is central to any concept of a self. In the next step, the transition is made from conceiving of oneself as an agent, which involves causally relating events in the world to one's actions, to then recognizing other subjects as goal-oriented agents, able to bring about changes in the world. The abstraction mechanisms described cannot account for all kinds of abstract cognition, such as mathematical-logical reasoning or composing symphonies. Nevertheless, this model can plausibly account for various cognitive abilities that are generally held to be rather complex and can thus be grounded in action-related representations. They are thus one the most fundamental kinds of representations, as they are involved in crucial cognitive functions at the heart of the behavior of a wide range of organisms, connecting humans with chimpanzees, squirrels and desert ants.

2 Being in and Toward the World: Body Schema and Motor Intentionality

Maurice Merleau-Ponty's 'Phenomenology of Perception' (1945/2012) is one of the first accounts to systematically address the role of the body in human cognition in general and in perception in particular. Central to Merleau-Ponty's notion of embodiment are the concepts 'body schema' and 'motor intentionality'.[4] Both concepts are of special importance for the development of an account of action-related representation. The two central aspects of action-related representation are the representation of features of the world relevant for interaction, and the subject's physical constitution. These two aspects are reflected, to some degree, in Merleau-Ponty's concepts of the 'body schema' and 'motor intentionality': Merleau-Ponty refers to motor intentionality as the subject's method of relating to the world, and refers to body schema as subjective information about the agent's body that is always related to events in the world (posture, skills, etc.). These two aspects, among other facets of embodiment, constitute the original, fundamental subject-world relation. Merleau-Ponty is relevant to the present discussion because he argues for a notion of intentionality (in opposition to Brentano (1874) that focuses on the body: Intentionality is not merely an intellectual characteristic of mental states, but is located primarily in the interacting body. All other aspects of intentionality are considered derivative to this original body intentionality. Another important aspect is Merleau-Ponty's emphasis that it is goal-directed action and not the mere movement that is foundational for cognition and perception. This aspect will be reflected in chapter 8, where a

4 Merleau-Ponty uses the terms 'motor project' and 'motor intentionality' interchangeably (Merleau-Ponty 1945/2012, S.113).

general account of action-related representation is presented. For these strong connections to the more recent and contemporary debates about action-related representation, as will be discussed in the following chapters, the careful analysis of Merleau-Ponty's ideas will be the point of departure and inspiration for developing an account of action-related representation that meets the requirements of philosophical analysis and empirical psychological evidence. Merleau-Ponty was always very eager to combine philosophical reasoning with data from empirical research, which is also the aspiration for this discussion.

2.1 The Body Schema

Maurice Merleau-Ponty's central aim was to overcome the pitfalls of empiricism and 'intellectualism' by assigning a central role for the body in the process of perception. Perception, according to Merleau-Ponty (1945/2012), is not an intellectual process. It is neither a product of the faculty of thought nor solely semantic content of the unity of consciousness, nor is it visual representation as the empiricist tradition would have it. Perception is above all a bodily process. Taylor Carman (1999) writes:

> Merleau-Ponty bases his entire phenomenological project on an account of bodily intentionality and the challenge it poses to any adequate concept of mind. [...] More generally, the problem of embodiment raises question concerning the very notion of the mental as a distinct phenomenal region mediating our intentional orientation in the world. Merleau-Ponty never doubts or denies the existence of mental phenomena, [...] but he insists [...] that thought and sensation as such occur only against a background of perceptual activity that we always already understand in bodily terms. (Carman 1999, 206)

It is thus the body which enables us to perceive; we are in the world through our lived body and the body is the medium and condition for perception. A psychology that focuses solely on mental representations or the experienced content of consciousness, while treating movement and action only as bodily processes obeying the commands of the thinking consciousness, would be invalid and incomplete.

The body schema is described by Merleau-Ponty as the locus of (implicitly) stored information about body parts and their position:

> I hold my body as an indivisible possession and I know the position of each of my limbs through a body schema […] that envelops them all. (Merleau-Ponty 1945/2012, 100)

The body schema thus consists of constantly updated information about the position of body parts, so that goal-oriented movement can be generated with respect to the current position of the individual parts. A distinction has to be made between the body schema and the so-called 'body image' (cf. Gallagher, 2001; 2005). The body schema is supposed to consist of unconscious information that is never explicitly represented, whereas the body image is meant to be conscious perception and thinking with the body as intentional object – looking at down my body provides me with a conscious, explicit representation of my body from the chest downwards. It is not entirely clear from Merleau-Ponty's writing whether he always considers the body schema to be unconscious information in principle, which can become conscious at times – turning into what is called a body image, or if the body schema is to be understood as an image of the body's posture that is generally available to conscious experience. Does the body schema consist of implicit knowledge or is it mainly explicit knowledge, implying that the posture of the body and the limb position would be explicitly represented in conscious experience? According to Merleau-Ponty, the body schema "was at first understood to be a *summary* of our bodily experience" and "thought to develop gradually throughout childhood and to the extent that tactile, kinesthetic, and articular contents associated between themselves or with visual content" (Merleau-Ponty 1945/2012, 101), being *the center of images.* This traditional understanding of the body schema supports an imagistic conception. In another passage, he states

> that the body schema is not merely an experience of my body, but rather an experience of my body in the world, and that it gives a motor sense to the verbal instructions. (Merleau-Ponty 1945/2012, 142)

Although not mentioning an imagistic conception here, this passage still suggests a 'conscious-experience-view' interpretation of the body schema. Elsewhere, Merleau-Ponty states that the body schema cannot be restricted to an association of images, suggesting a conception of the body schema that is law-like, resembling a plan:

> Rather, these associations must be constantly submitted to a unique law, the spatiality of the body must descend from the whole to the parts, my left hand and its position must be implicated in an overall body *plan* and must have their origin there [...] (Merleau-Ponty 1945/2012, 101).

He goes on to suggest a second definition of the body schema, which should provide more clarity:

> [...] it will no longer be the mere result of association established in the course of experiences, but rather the global awareness of my posture in the inter-sensory world, a "form" in the Gestalt psychology's sense of the word. (Merleau-Ponty 1945/2012, 102)

By global awareness, Merleau-Ponty refers to information about the body, such as its posture, which is poised for further processing. It is information of the body as a whole, not just the individual parts, that gets constantly updated. The body is always situated, so the body schema also contains information about the body's posture in relation to the surrounding environment and the objects therein. The passage that most clearly reveals that Merleau-Ponty thinks the body schema cannot be confined to the consciously experiential body is the following:

> If the need was felt to introduce the new word [the body schema; T.S.], it was in order to express that the spatial and temporal unity, the inter-sensorial unity, or the sensorimotor unity of the body is, so to speak, an in principle unity, to express that this unity is not limited to contents actually and fortuitously associated in the course of our experience, that it somehow precedes them and in fact makes their association possible. (Merleau-Ponty 1945/2012, 102)

Here, it becomes obvious that Merleau-Ponty conceives of the body schema as being constitutive for experience and makes a distinction between information about the sensorimotor unity of the body and conscious experience of one's own body. The function of the body schema is described as effectively enabling interaction with objects in the world and through this interaction, providing a sense of being "in and towards the world" (Merleau-Ponty 1945/2012, 103). Every object perceived by a subject is perceived as a figure standing out against a background, and this relation is perceived in relation to one's own body. So every perception of an object involves perceiving the body (cf. Merleau-Ponty 1945/2012, 103). The body schema thus plays a constitutive role for object representation as it is always implied in every perception of an object and at the same time an expression of the interaction possibilities perceived in accordance with these objects. The body schema as such is for action and always represents the body and its parts in their situatedness towards objects. In order for a subject to grasp a perceived object front of her, she must know where the object is in relation to her arm and hand positions.[5] Bodily space and external space form a practical system, the system being constitutive for objects actually becoming a part of an action-goal, and thus it is in action that bodily space manifests itself (cf. Merleau-Ponty 1945/2012, 105). This conditional relation of body space and action is reflected in the debate about egocentric and allocentric spatial representations, the former being described as being representations that are already representing in an action format (cf. Vosgerau 2009, ch. 7.2.3; more on egocentric representation and their role in action will follow in ch. 4 and ch. 8 of this book).

5 The whole posture will be represented in the body schema, but the aspects mentioned (arm and hands) are the most important ones for a concrete grasping action and will therefore be more salient in experience than e.g. the legs positions: "If I stand in front of my desk, and lean on it with both hands, only my hands are accentuated and my whole body trails behind them like a comet's tail. I am not unaware of the location of my shoulders or my waist; rather, this awareness is enveloped in my hands and my entire stance is read, so to speak, in how my hands lean upon the desk." (Merleau-Ponty 1945/2012, 102)

In representing body space as action space, Merleau-Ponty uses the notion of the body schema for constituting a kind of subjective, action-related knowledge that corresponds to a mode of being[6]: In a study cited by Merleau-Ponty, a patient was unable to point to a part of his body unless he was also instructed to grasp it (cf. Merleau-Ponty 1945/2012, 106). The patient thus could only perform the relevant movements if they were part of an action including the anticipation of a location as action goal. A purely descriptive pointing or otherwise unmotivated movement was unable to be exercised by the patient. Merleau-Ponty concludes from that, that there are different ways to have knowledge of a location (cf. Merleau-Ponty 1945/2012, 106). He introduces two kinds of knowledge: a practical knowledge, underlying actions, and a more descriptive knowledge, specifying spatial locations in an objective sense. The patient seems to have access only to the former kind of knowledge, where a location is presented as a goal state of a grasping action. The location in the patient's case, his nose, is part of a bodily knowledge when it comes to performing an action, the arm and hand "knows" where to find the nose when intending to grasp it, but there is no equivalent knowledge when the patient should just point to the nose – which implies having a more detached knowledge where the objective location of the nose is. How does Merleau-Ponty explain this difference between the abilities of healthy subjects and the patient? The subject executing habitual, familiar actions and action patterns does not need to represent her body as something with objective spatial properties, thus subject does not represent the body as an object among others which she could simply designate by, e.g., pointing. The subject in these kind of situations is not even aware of the movements she needs to generate, the adequate movements are elicited because the subject is part of a body-world system, in which the body-object relation immediately and implicitly (i.e., not objectively represented) determines the action possibilities and the movements required. This is strikingly similar to Gibson's (1986)

6 "The two "stimuli" are only genuinely distinguished if we take into consideration their affective value or their biological sense; the two responses only cease to merge if *Zeigen* and *Greifen* are considered as two different ways of relating to the object and two types of being in the world." (Merleau-Ponty 1945/2012, 124)

idea of affordance perception (for a detailed discussion of Gibsonian affordances, see ch. 3).[7]

The following passage shows how Merleau-Ponty is actually anticipating the idea of affordances as action-related properties and the idea of action-related representation of one's environment in general. It is a rather long quote, but worth reciting, as it captures the essence of Merleau-Ponty's account of action-related experience:

> Between the hand as a power for scratching and the point of the bite as a place to be scratched, a lived relation is given in the natural system of one's own body. The operation takes place wholly within the order of the phenomenal, it does not pass through the objective world. […] Likewise, the subject placed in front of his scissors, his needle, and his familiar tasks has no need to look for his hands or his fingers, for they are not objects to be found in objective space (like bones, muscles, and nerves), but rather *powers that are already mobilized by the perception of the scissors or the needle*, they are the center-point of the "intentional threads" that link him to the given objects. We never move our objective body, we move our phenomenal body, and we do so without mystery, since it is our body as a power of various regions of the world that already rises up toward the objects to grasp and perceive them. Likewise, the patient need not seek a situation and a space in which to deploy concrete movements, this space is itself given, it is the present world*: the piece of leather "to be cut" and the lining "to be sewn." The workbench, the scissors, and the pieces of leather are presented to the subject as poles of action;* […] that calls for a certain […] labor. The body is but one element in the *system* of the subject and his world, and the task obtains the necessary movements from him through a sort of distant attraction, just as the phenomenal forces at work in my visual field obtain from me, without any calculation, the motor reactions that will establish between those forces the optimum equilibrium […]. (Merleau-Ponty 1945/2012, 108f, my italics)

The world, according to the passage quoted, is phenomenally presented to the subject in terms of possible actions. It is the bodily space of the

7 As Gibson never mentions Merleau-Ponty, it is unclear if there was a direct influence at all. However, both Gibson and Merleau-Ponty have striking parallels in their work, which is most likely due to their *gestalt* background which heavily influenced their research. (cf. Sanders 1993)

subject that determines her action space, and this relation is a fundamental one. The patient has problems conceiving of his body as something detached, as something that can be localized within objective space, but has no problems conceiving of the world as a space of possible interaction. Of course, it is debatable if Merleau-Ponty's distinction between a *Greifen* and a *Zeigen* can be treated in such a different way, especially as *Zeigen* is clearly an action. What is an important insight though is that the goal of an action, the endpoint matters for the way how we represent actions and the action space: The nose in the example of the patient is not a thing with a defined, objective location, but just the endpoint of a grasping movement, the nose is represented in terms of this very movement. The point to point at is represented as a point within objective spatial coordinates of which the body is just another space-point, and thus the patient fails to localize it – he lost the ability to represent his body in a detached view, as an image. The use of the expressions in inverted commas, "the piece of leather 'to be cut' and the lining 'to be sewn'" sounds very similar the idea of Gibsonian affordances. The piece of leather has action-relevant properties, which Gibson would phrase as "being cut-able", or, if put in representationalist terms, the subject representing the leather as something associated with the action "to cut", in terms of a possible cutting action. The objects in a subject's world, according to Merleau-Ponty, are hence not perceived an objective, detached way, e.g. their shape and color properties, but the very object is transparent, i.e. what is perceived is its meaning for possible actions.

This passage of 'The Phenomenology of Perception' is clearly an early precursor to the idea of action-related representation, relating perceived/represented action possibilities of objects to the perceiving subject's body. The relation in question is, following Merleau-Ponty's concept of the body schema, a constitutive one: it is via the body, and especially the bodily information residing in the body schema, that the subject perceives action possibilities. The body schema thus is an integral part of action-related representation in Merleau-Ponty's account and will be central to the general account of action related representation, as will be presented in chapter 8.

2.2 Motor Intentionality

The other central notion in Merleau-Ponty's account of embodied perception is the idea of motor intentionality. In the long quote on the previous page, parts of the idea of motor intentionality are already presented by referring to the "intentional threads" which link subjects to objects in the world. Motor intentionality refers to the object directedness of every action, which is given in terms of possible movements towards the object. It is not a process of conscious thought, though: a subject might be able to (intellectually) understand motor instructions, but nevertheless be unable to transform them into the appropriate movements, though the subject is capable of executing the movements in principle. For Merleau-Ponty, this finding leads to the conclusion that there is a capacity, a mode of the body that consists in "an anticipation or a grasp of the result assured by the body itself as motor power, a 'motor project' (*Bewegungsentwurf*), or a 'motor intentionality' without which the instructions would remain empty" (Merleau-Ponty 1945/2012, 113).

Motor intentionality is the body's understanding of its environment, a way of grasping the environment in motoric ways. Motor intentionality, as conceived by Merleau-Ponty, is fundamentally related to the body's knowledge of spatial relations. The body knows where its limbs are, so in order to reach for my knee, I do not have to think or search for it, but just reach for it. Generally, Merleau-Ponty seems to think that the directedness or aboutness of the motor intentionality is an incorporation of objective space "into [the subjects] bodily space" (Merleau-Ponty 1945/2012, 146).

However, motor intentionality is not just a spatial relation, or a mode of conceiving of spatial relations: Merleau-Ponty also emphasizes the role of the object in one's perceived action space. Nevertheless, Merleau-Ponty's idea of the embodied subject is mainly a spatial notion of embodiment: the motor space, or the action space as a mode of intentional relation to the world. It is the actual grasp of objects in the subject's action space that is enabled by the body's motor intentional access to the world, which is a mode of the embodied grasp of the world. But something is still missing in Merleau-Ponty's idea of motor intentionality, and this missing bit appears to be crucial: As Kelly (2002) points out, one can point to a location without being directed at a specific object as well as it if it would involve an object, but one cannot perform[8] a grasping movement in the same manner without it being object directed. Hence, the location of the object and its location in the subject's motor space can only be one aspect of the subject's motor intentionality. Kelly seeks to overcome this shortcoming of Merleau-Ponty's by claiming that a grasping action, in contrary to a pointing action involves the "entire object, not just [...] some independently specifiable spatial feature of it" (Kelly 2002, 384), which is supported by Merleau-Ponty's claim that we reach out for *specific things*:

> The gesture of reaching one's hand out toward an object contains a reference to the object, not as a representation, but as this highly determinate thing toward which we are thrown, next to which we are through anticipation, and which we haunt. Consciousness is being toward the thing through the intermediary of the body. (Merleau-Ponty 1945/2012, 140)

The dimension of the object as such and not *just* its location is acknowledged by Merleau-Ponty, but he does not provide any further characterization of how the subject anticipates the object. And this is where the object's action related properties, its affordances, enter the stage. By referring to the object's affordances, such as its size for grasping, its weight for picking up, its surface and temperate etc., one could flesh out how the subject anticipates the object. There is no need to perceive or relate to the entire object, whatever that might mean anyway, as long as the properties essential for the intended interaction are perceived, referred to or even represented, the subject will be able to perform a grasping action directed at the object and thus express its motor intentionality in terms of its body schematic information and the anticipation of the location and affordances of the object. Merleau-Ponty does not consider the action related properties to any greater detail, for him it is enough to show that the subject's approach to the world is an embodied one, which manifests itself in the action orientation of the subject. This action oriented approach is

8 At least this grasping movement without being directed at an object will not be performed with the same precision, or, with Merleau-Ponty's words: "From its very beginnings, the grasping movement is magically complete; it only gets under way by anticipating its goal [...]" (Merleau-Ponty 1945/2012, 106).

more fundamental than any cognitive approach that relates subjects to their environment via cognitive processes of reflective thought. Subjects perceive the world through their body and because they are embodied, and sophisticated cognition is constituted by the subject's body.

It is exactly Gibson's (1986) enterprise of establishing an ecological psychology that attempts to shift the focus to the subject's environment and the action-related properties one detects therein – while maintaining Merleau-Ponty's basic ideas on the role of the body schema and motor intentionality for perception and cognition.

3 Perceiving Possible Actions: Gibson's Affordances

3.1 Introduction

This chapter will critically examine Gibson's concept of affordances (Gibson 1966; 1986). The concept of 'affordances' being a central notion in the field of ecological psychology is also widely used in the cognitive sciences, empirical psychology and philosophy and plays a prominent role in art and design. In spite of, or maybe due to its wide and almost commonsensical use, the notion of affordances up to the present day has remained a very controversial term, whose nature is either hotly debated or left unspecified in many cases. One of the main problems with the terms affordances is thus that there exist a number of definitions or unspecified uses that eventually lead to a lot of confusion about the nature of affordances and its value in scientific discourse. Although explicitly invented to be a non-representational account of perceiving action possibilities in the environment, it has to be a part of the discussion of action-related representation. The reason is obvious: Gibson's entire focus was on the interaction possibilities the environment offers and how animals perceive these possibilities. Thus, the whole enterprise of his later works was to account for the perception of possibilities for interaction while establishing an alternative to empiricist and representational models of perception and cognition. In this sense, Gibson offered a non-representational account of action-related representation.

The aim of discussing Gibson's notion of affordances is to show that his work is addressing the right problems and is, in its radical focus on interaction with the animal's environment, providing an inspiring perspective on the problem of accounting for the cognitive processes enabling goal directed actions to subjects of all species.

A further aim of this chapter is also to show why and how Gibson's approach is severely flawed in important respects and can therefore not establish a radical change in how psychology should think about the perception of action possibilities. His concept of affordances though, reformulated in a representational spirit, can be made compatible with the (mainly) representational research paradigm of contemporary cognitive science. Affordances, understood in a representational way, can still do a lot of explanatory work, especially when it comes to explaining, for example, development of simple affordances perceived by toddlers towards the representation of affordances for other agents and the cognitive processes underlying these abilities. This chapter is therefore a criticism of Gibson's anti-representational view on affordance perception and a reinterpretation of the concept of affordances in representational terms to save the explanatory potential and provide a more coherent concept that can be applied in cognitive research.

The term 'affordances' was invented by J. J. Gibson (1986) and has mainly been dealt with, until recently, in the area of ecological psychology. It has been mostly advocates of the general ecological psychology agenda who were concerned with providing theoretical justification of Gibson's central ideas. This is especially true for the two most controversial and revolutionary ideas, the possibility and necessity of direct perception and – strongly connected to direct perception – the notion of affordances. This is problematic or difficult because ecological psychology is a rather idiosyncratic enterprise, involving the (radical) departure of many beliefs and theoretical presuppositions held by more traditional approaches to psychology, especially those with a "cognitive approach". Trying to comprehend the concept of affordances without considering its origins and backgrounds in ecological psychology can only be a partial approximation. The first part of this chapter will therefore try to reconstruct the notion of affordances and present the most important theoretical reflections mainly from the field of ecological psychology, starting with Gibson's own proposal and his theoretical background of ecological psychology. There are few theorists who discuss and try to provide an account of affordances that do not descend into the field of ecological psychology or at least accept a lot of its premises. The main premise the ecological faction shares and which distinguishes them from most other, non-ecologically branded philosophers and psychologists also using the term affordances, is their anti-representationalism. From this, it follows that the first approach to a discussion of the concept of affordances is clarifying if affordances in an anti-representationalist framework means and implies the same as in the other frameworks, which are in most cases explicitly or implicitly representationalist. Basically all areas of psychology and related disciplines refer to affordances when they want to refer to functional properties of objects or interaction possibilities for agents. Due to this frequent underspecified use it is often quite hard to decide what exactly is referred to by the term affordances: Is it possibilities for (inter-)action, meaning situational features; is it properties of objects, such as handles, grips, lids or openings; or does it refer rather to action capabilities of subjects such as bodily constitution, physical skills and abilities, which enable subjects to interact with objects in their environments?

This chapter aims at finding a definition of affordances, which avoids the problems that arise out of commitment to the presuppositions of ecological psychology and at the same time being a substantial scientific concept nevertheless. By arguing and providing evidence for the claim that affordances are action-related representations, which explicitly represent properties of objects relevant for interaction in accordance to the physical condition of subjects, the concept of affordances will be made applicable for cognitive sciences accepting representations as fundamental for cognition. Moreover, affordances defined as action-related representations can provide important insights in basic cognitive abstraction mechanisms.

In the following sections I will give an overview of Gibson's ecological approach to visual perception in general, laying the foundations for a better understanding of Gibson's theory of affordances. I will thus begin with Gibson's central claim that perception is direct and substantially different from what perceptual theorists held to be true. It is based on an anti-representationalist view of perception which argues for perceptual experience being unmediated by mental states and therefore consisting of the act of information pickup.

The information most relevant to pick up for animals is the information specifying affordances of the environment. Hence, affordances are specified directly in the ambient light array, which means that meaning and values are already to be found in the structured information given by the light arrays. Understanding Gibson's original proposal of the theory of affordances is important for the subsequent interpretations of the notion of affordances.

3.2 Short Introduction to Gibson's Theory of Perception

Gibson's theory of affordances can only be understood and analyzed in the light of his theory of (direct) perception to which Gibson also referred to as "theory of information pickup" (cf. Gibson 1966, ch. 13; Gibson 1986, ch. 14). The theory of information pickup states that organisms forming perceptual systems are surrounded by "available stimulation […][that] has structure, both simultaneous and successive and that this structure depends on sources in the outer environment" (Gibson 1966, 267). Hence perception consists in registering "the invariants of this structure" (Gibson 1966, 267) and "meaningful information can be said to exist inside the nervous system as well as outside". (Gibson 1966, 267) This forms the core of Gibson's idea of direct perception, i.e., perception which is not mediated by mental states:

> The brain is relieved of the necessity of constructing such information by any process – innate rational powers (theoretical nativism) the storehouse of memory (empiricism), or form-fields (Gestalt theory). The brain can be treated as the highest of several centers of the nervous system governing the perceptual systems. Instead of postulating that the brain constructs information from the input of a sensory nerve, we can suppose that the centers of the nervous system, including the brain, resonate to information. (Gibson 1966, 267)

According to Gibson, no mental states, such as mental representations or memory states, are the mediators or bearers of the perceptual experience, but the information that is available is already structured in the environment, and is registered by the perceptual system by means of resonating to informational invariants and variations. Gibson's view on perception stems from his criticism of "the orthodox theory of the retinal image" (Gibson 1986, 58) and his general criticism of the so-called sensation-based theories of perception. Gibson has two main objections, one which is based on an alleged fallacy he calls "'the little man in the brain' theory" (Gibson 1986, 60), the other a general objection against the idea that brain states could sensibly represent the qualities represented by the retinal stimuli. Let's have a closer look on both objections.

The theory of perception Gibson is referring to originates in Johannes Kepler's theory of image formation which states that light "forms an image of an object on the back of the eye" (Gibson 1986, 58). The image of the object is formed by a multitude of "focus points" on the back of the eye, directed there by the lens which bundles rays of light. Every point of the object emits an infinite number of light rays, of which some get bundled by the lens and focused on a single point on the back of the eye. Every radiation point corresponds to a focal point and the sum of focal points assembles the image of the object. According to Gibson this "was and still is the unchallenged foundation of the theory of image formation" (Gibson 1986, 59). This model of vision might work well and has proven successful e.g. in camera building, where an image of an object is literally projected on a screen-like surface, but it is misleading when it comes to vision and perception. Thinking of vision this way would require a perceiver of the retinal image – the little man in the brain, a homunculus who is actively looking at the retinal image. This in turn would of course imply that the homunculus had eyes himself and a retinal projection and thus lead to an infinite, paradoxical iteration.

The second line of criticism addresses a version of the theory of image formation – the sensation-based theories of perception. According to Gibson's interpretation of the sensation-based theories of perception, the "correspondence between the spots of light on the retina and spots of sensation in the brain can only be a correspondence of intensity to brightness and of wavelength of color." (Gibson 1986, 61) Gibson is doubting that this can be enough or the right kind of information for full-blown perception of the environment:

> If so, the brain is faced with the tremendous task of constructing a phenomenal environment out of spots differing in brightness and color. If these are what is seen directly, what is given for perception, if these are the data of sense, then the fact of perception is almost miraculous. (Gibson 1986, 61)

Retinal stimulation cannot be the right informational source for perception as this is too poor a stimulus to be the cause of the rich perceptual experiences animals and humans have. Therefore, the informationally rich environment itself has to be the source and the (unmediated) cause of perceptual experience. Moreover, Gibson stills sees necessity for "a little man in the brain", even if there is no analog pictorial projection but more of a digital data transmission from retinal stimulation to brain activation, as these signals have to be sent in a certain format – in a code – and be decoded or interpreted afterwards. This would again lead to a homunculus-like picture of the mind as a subject interpreting sense data and thus be prone to the same criticism as outlined above. Gibson generally rejects the idea of information as signals or codes that have to be encoded by the perceiver, as this is the erroneous consequence of a fallacious view of information:

> We tend to think of information primarily as being sent and received, and we assume that some intermediate kind of transmission has to occur, a 'medium' of communication or a 'channel' along which the information is said to flow. Information in this sense consists of messages, signs, and signals. [...] The ambient stimulus information available in the sea of energy around us is quite different. The information for perception is not transmitted, does not consist of signals, and does not entail a sender and a receiver. The environment does not communicate with the observers who inhabit it. Why should the world speak to us? The concept of stimuli as signals to be interpreted implies some such nonsense as a world-soul trying to get through to us. The world is specified in the structure of the light that reaches us, but it is entirely up to us to perceive it. The secrets of nature are not to be understood by the breaking of its code. (Gibson 1986, 62f)

Gibson's proposition to overcome these problems of conventional theories of optical information and visual perception is his "theory of information pickup" (Gibson 1986, 238). Central to this idea is that perception is not an interpretation of sense data or stimuli delivered by the senses in isolation, but that perception occurs only and necessarily in "perceptual systems" (Gibson 1986, 244). Perceptual systems have a number of qualities that distinguish them from mere senses, which are defined as "a bank of receptors or receptive units that are connected with a so-called projection center in the brain" (Gibson 1986, 245). Thus, perceptual systems are more than receptive cell units. They comprise of all the (bodily) parts involved in perceptual events or processes. In visual perception, this means not only the eyes and the visual cortex, but the moving head and the rest of the body that can be adjusted to create new visual stimuli, e.g. by turning the head or body or changing the position. Information for perception is therefore obtained actively, whereas the senses are considered to be passive receptors. Information through the senses can only be recombined and associated, information of the perceptual system can be learned – by this, Gibson means that perception is itself a process of learning and development, subjects actively have to learn to perceive, which is a lifelong process and can be more or less "subtle, elaborate and precise" (Gibson, 1986, 245). Perception in that respect is an active skill, and the pickup of information can be better or worse across subjects. This is the strategy by which Gibson attempts to counter objections that his theory of direct perception does not allow for misrepresentation or non-veridical perception, which would commit Gibson to the strongest form of realism possible. I shall return to this point later.

Another critical aspect of perceptual systems is that perceptual systems react to the qualities things have in the environment, these qualities mainly being what these things afford, whereas special senses have receptoral stimulation as inputs (cf. Gibson 1986, 246). This is also one of the main arguments for Gibson's theory of direct perception: If the basic units of perception would be stimulations of the sensory receptors and then signals conveyed by them to cortical areas for further processing, the perceiver would be cut off from the world, because objects in the environment would be in no respect similar to the outputs or patterns of receptoral stimulations. The external cause of these stimulations cannot be easily deduced from these stimulations and therefore we cannot gather knowledge about the exterior world or, even more problematic, we would already have to know what to perceive in the external world in order to correctly interpret the stimuli. The only way out of this alleged dilemma for Gibson is to claim that perception is directly about the qualities (the affordances) of the environment and stimulations of receptors do not play an elementary role, and suggests a direct relation of the whole perceptual system to the qualities of the external objects:

> The alternative is to assume that sensations triggered by light, sound, pressure, and chemicals are merely incidental, that information is available to a perceptual system, and that the qualities of the world in relation to the needs of the observer are experienced directly. (Gibson 1986, 246)

The next important feature of the theory of information pickup is the postulate of a constant flow of information and the rejection of the notion of discrete, analyzable stimuli sequences. The flow of information in the ambient light array specifies the qualities of the surrounding sufficiently, so that the observer has only to direct his attention to the invariant structures in the ambient light array. What is available for the perceiver is information for persistence and change, of both the perceiver (e.g. self-motion) and the objects in the environment. Traditional theories of sense-data-like perception have to assume that the perception change and persistence is the outcome of a comparison of two sense data whereas Gibson's invariant structures are themselves specifying change or persistence – no mental comparison is needed according to this notion. This leads to Gibson's way of securing the correct perception of identity of objects and persons with the theory of information pickup, that by definition cannot rely on comparison of actual stimuli and stored stimuli or representation in memory, as the traditional approach according to Gibson would have it:

> In the case of the persisting thing, I suggest, the perceptual system simply extracts the invariants from the flowing array; it resonates

> to the invariant structure or is attuned to it. In the case of substantially distinct things, I venture, the perceptual system must abstract the invariants. The former process seems to be simpler than the latter, more nearly automatic. The latter process has been interpreted to imply an intellectual act of lifting out something that is mental from a collection of objects that are physical, of forming an abstract concept from concrete percepts, but that is very dubious. Abstraction is invariance detection across objects. But the invariant is only a similarity, not a persistence. (Gibson 1986, 249)

In this sense, objects and person have features that are invariant to a certain extent. This detection or abstraction of these very invariances, or the resonation is what is traditionally understood as identification of persons or objects.

Gibson conceives of perception as an active process, the active attunement to information, which does not or cannot be stored in memory or be transferred from a sender to a receiver, but has to be attended to. Attending to information is the same as information pickup in Gibson's terminology; the information does not have to be stored in memory (and then retrieved, compared, associated etc.) because the information is *always available* in the ambient optical light array structured by the features the environment actually exemplifies. That said, Gibson's theory of information pickup is a radical externalist theory of perception, as the content of perception is external to the perceiver, being already specified and always available in the energetic structures surrounding a perceptual system. Gibson is also specific on what is and what is not perceived in the light of information pickup:

> places, attached objects, objects, and substances are what are mainly perceived, together with events, which are changes of these things. To see these things is to perceive what they afford. This is very different from the accepted categories of what there is to perceive as described in the textbooks. Color, form, location, space, time, and motion-these are the chapter headings that have been handed down through the centuries, but they are not what is perceived. (Gibson 1986, 240)

Perceiving the aforementioned entities is to perceive what they are for, their function and role for possible interactions. Of course, this is a clear departure from traditional accounts of perception, which are mainly concerned with primary and secondary properties. This will become more explicit in the next chapter about Gibson's notion of affordances.

To conclude this overview of Gibson's theory of information pickup, his understanding of the relation between perception and knowledge has to be briefly mentioned. Gibson proposes a new approach to knowing provided by the theory of information pickup, one that "makes a clear-cut separation between perception and fantasy, but [...] closes the supposed gap between perception and knowledge" (Gibson 1986, 258). He defines perception and knowing to be basically the same things in that the same processes are underlying both. There is a only a difference in degree, but not in type of process. The very same processes and systems that enable perceivers to perceive the world is providing knowledge about the world, as "[k]nowing is an *extension* of perceiving." (Gibson 1986, 258) This again is based on the assumption that the process of extracting and abstracting invariant structures in the ambient energy flux not only enables perceptual awareness but at the same time constitutes knowledge. This can only be secured by perception on the basis of the detection of invariant structures of the environment being always and necessarily veridical (cases of misperception undergo a special treatment, cf. ch. 6.2.3 of this book).

Only if perception is already veridical, it can be extended to knowledge proper, with perception being "the simplest and best kind of knowing" (Gibson 1986, 263). Moreover, this implies that if perception is direct in Gibson's sense of being unmediated by anything mental (images, cognitive processes, representations etc.), the same holds for knowledge which also has to be direct because it should be of the same kind. Gibson does allow for mediated forms of knowledge, the most common being mediated by instruments (magnifying glasses, telescopes), by (verbal) descriptions or by pictures. All these derived forms of knowledge extend perception further, but are conceived to be still in one line with perception, not being different in type.

To sum up, the main features of Gibson's theory of information pickup can be described as:

- Perception is not based on special senses, input of signals or stimulation of receptors but on perceptual systems resonating to the ambient energy flux.
- The information for perception is already structured in the ambient energy array and not signals that have to be interpreted.
- The activity of the perceptual system consists of the detection, extraction and abstraction of invariant structures in the ambient energy flux, persistence and change being the crucial features to be specified by invariant structures.
- Perception is direct and mainly affordances of the environment are what is perceived, not the traditional qualities such as form, color, motion etc.
- Perception and knowledge are continuous processes and in principle the same, therefore both being direct.

3.3 Gibson's Affordances: Relational Animal-Environment Properties

Of all the possible features that can be specified in the ambient energy flux and therefore be directly picked up, affordances are the most important and at the core of Gibson's ecological psychology enterprise. By seeing places, events, surfaces, objects, etc., the observer picks up information regarding what they afford – what can be done, the functional specification of the environment. The famous, brief definition of affordances Gibson initially gives has become a standard paraphrase in the literature:

> The *affordances* of the environment are what it *offers* the animal, what it *provides* or *furnishes*, either for good or ill. The verb *to afford* is found in the dictionary, but the noun *affordance* is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment. (Gibson 1986, 127)

According to this definition, affordances are functional features of the environment and have to have a certain value for the animals, either positive or negative. Affordances are complementary, which means they are features of the environment relative to animals. Thus, the surface of the environment affords support for some animals, relative to the weight of the animal. Water in this respect can afford support for some species, not for others. Surfaces can also afford a number of different things: "Terrestrial surfaces, of course, are also climb-on-able or fall-of-able or get-underneath-able or bump-into-able relative to the animal." (Gibson 1986, 128) This list of possible affordances of the environment can be arbitrarily continued: different objects are sit-on-able for humans (relative to knee-height), water affords being drink-able, a pathway affords being walk-through-able for some species of animals. Affordances in this sense offer various behavioral possibilities relative to the physical conditions (and skillfulness) of animals. For Gibson, these behavioral possibilities equal "values and meanings" of the environment and the objects therein. What is special about Gibsonian affordances is that they are of an objective nature, although Gibson remains rather ambiguous about this:

> An important fact about the affordances of the environment is that they are in a sense objective, real, and physical, unlike values and meanings, which are often supposed to be subjective, phenomenal, and mental. But, actually, an affordance is neither an objective property nor a subjective property; or it is both, if you like. […] It is equally a fact of the environment and a fact of behavior. It is both physical and psychical, yet neither. (Gibson 1986, 129)

They are subjective and objective in nature at the same time, or might even be understood as a third class of properties that exceeds the dichotomy of subjective and objective. Furthermore, he claims that affordances are directly perceivable:

> to perceive them (the composition and layout of surfaces] is to perceive what they afford. This is a radical hypothesis, for it implies that the "values" and "meanings" of things in the environment can be directly perceived. Moreover, it would explain the sense in which values and meanings are external to the perceiver. (Gibson 1986, 127).

As became evident in the preceding chapter, an important aspect in Gibson's theory is that perception consists of proprioception and exteroception. This concept of perception is of major importance for Gibson's view on affordances: every act of perception of external values and meanings implies perceiving information about the perceiver's body simultaneously:

> This is only to reemphasize that exteroception is accompanied by proprioception – that to perceive the world is to perceive oneself. (Gibson 1986, 141)

Gibson's claim that there are features of the environment that allow for interaction, which would be labelled in more contemporary terminology as action-related properties or functional properties, is widely acknowledged. There is a common sense concerning the wide, rather unspecified use of the term affordances and synonymous terms. Also, the idea that functional features of the environment are related to the physical condition of animals is not shocking: it seems quite reasonable to assume that objects are only graspable for creatures with hands (or something functionally equivalent) of the appropriate size. The controversy about Gibson's notion of affordances is that he strongly insists that affordances are different from everything physics told us about physical properties and perception thereof.[9] Affordances are directly perceivable, and as meanings and values are the same as affordances, meanings and values, qua being external to the perceiver, are directly perceivable too. This results in Gibsonian affordances having two peculiar and controversial characteristics:

1. Affordances are a different kind of property from what we normally take to be physical properties (subjective and objective, or both; neither physical nor phenomenal)
2. Affordance perception is direct, hence not "mediated", meaning they are neither represented nor inferred.

These two characteristics of affordances depend on each other to a certain degree: without affordances being objective in the sense that they really

9 Gibson opposes "physical physics" to "ecological physics" (cf. Gibson 1986, 139).

exist independently of being perceived, they could not be "picked up" directly in an unmediated way. This of course does not entail that the opposite holds: not everything that exists objectively, independently of a perceiving system, is already directly perceivable – if direct perception is possible at all. Being objective does not guarantee being directly perceivable. Affordances being external and the perception of affordances being "a process of perceiving a value-rich ecological object" (Gibson 1986, 140) on the other hand is necessary for the possibility of unmediated perception. If meanings and values would be added by the perceiver to a neutrally perceived object, this would be precisely the kind of inferential mental process Gibson strongly rejects in his theory of information pickup. These two features of affordances, special and controversial in their nature, require a closer examination of the way Gibson introduces, defines and justifies these claims.

First, let's have a look at how Gibson argues for the special nature of affordances as properties, which are neither physical nor phenomenal, both subjective and objective at the same time. "Affordances are properties taken with reference to the observer" (Gibson 1986, 143), Gibson writes, and refers to properties that cannot be specified without an observing system. The classical notion for relational properties is the distinction between primary and secondary qualities, primary qualities existing independently of being perceived, such as mass or shape, secondary qualities being dependent on the perceiver, such as colors and sounds. The theory of two qualities claims that the color of a given object can only be specified in terms of who perceives it, as phenomenal quality depending on the perceiver's abilities. A thing might be red for a human but have some sort of greyish-yellow shape for a cow, as the cow's color perception abilities are different to ours. On the contrary, an object will have the same shape no matter who perceives it or if there's a perceiver at all. The infamous sceptic's question, if a falling tree makes a sound if there is no one to listen, illustrates the idea of there being objective and subjective, phenomenal qualities. But this cannot be what Gibson has in mind when he stresses the relational nature of affordances, as he is quite clear in his dismissal of that very distinction:

> These seven modes or qualities take the place of the so-called modes of appearance of color [...]. And, when surface layout is also considered, they take the place of the so-called qualities of objects, color on the one hand and "form, size, position, solidity, duration, and motion" on the other. These latter are John Locke's "primary" qualities, those that were supposed to be "in the objects" instead of merely "in us". This distinction between primary and secondary qualities is quite unnecessary and is wholly rejected [...] (Gibson 1986, 31)

The main argument Gibson give for the rejection of the primary/secondary distinction is that subjective modes of appearance and representation cut the perceiving system off from its environment and would therefore open the doors for a radical epistemological skepticism, as the only thing perceivers can relate to and therefore perceive are sense data or something equivalent. And of course he rejects any theory that presupposes inner vision, as that would imply a homunculus, which would be paradoxical and lead to strange regresses. What he offers instead is more of a collection of (ecological) properties things can possibly have while taking them as directly perceivable. Gibson wants to substitute the traditional set of qualities including color, form, size, position, solidity, duration, motion etc., with an ecological description of the surfaces of things (substances). He speaks of luminous surfaces as distinguished from illuminated surfaces, of reflectance, smoothness, roughness, opacity and so on (cf. Gibson 1986, 31). By doing this, Gibson is purely focusing on the physical aspects of surfaces that emit and absorb various reflectance spectra and thus specify invariants in the ambient energy array. Coal, for instance, "has a low reflectance (about 5 percent), and snow has a high reflectance (about 80 percent)." (Gibson 1986, 30) The perceptual system resonates to the structure specified in the light array and thus perceives coal or snow. Gibson calls this *characteristic reflectance*, by which he means that every object and every substance uniquely specifies invariants in the ambient energy flux. Hence, the difference of qualities perceived should be explained with invariant structures that do not need a further distinction in qualities that are more or less dependent on perceiving systems. Accordingly, affordances are neither primary nor secondary qualities, as these do not exist in Gibson's view. Affordances thus are information structured in the ambient light array. Perceiving one's environment is perceiving its affordances, and therefore all information available about one's environment specifies at the same time its affordances.

It is not as easy as that, unfortunately. In addition, Gibson also emphasizes that affordances (values, meanings) are external to the perceiver. This means that things afford what they do because of their physical properties and the physical properties in turn are specified directly in the light. Every property, every object being a compound of properties has a unique way of structuring light. It is due to this one-to-one relation that information can be *picked up* and does not have to be interpreted or be the outcome of inference. Meanings and values are part of the external world and not part of the mental world of animals. This is what Gibson means when he claims that affordances are objective. But affordances should be subjective at the same time, and as meaning and values have to be external, the subjective aspect of affordances has to be found elsewhere. According to Gibson, an affordance "points two ways, to the environment and to the observer […][as] does the information to specify an affordance." (Gibson 1986, 141) Information about the environment goes along with information about the body of the perceiver, proprioception accompanies exteroception. It is not fully clear what exactly this statement entails. First, it is fair to state that for Gibson, the subjective aspect of affordances is assumed to be the self-perception of the perceiver that accompanies the perception of the external world. Second, this self-perception (proprioception) is assumed to be a part of the affordance – the information for an affordance is given in terms of object properties *and* subject properties. However, this makes it even more difficult to understand what Gibson means by affordances being external to the perceiver.

Either, affordances are only partly external, being also about object properties while at the same time internal, as it is only a particular subject perceiving an affordance according to his physical constitution. The perception of one's own body would then guide or limit the perception of affordances. This is one possible way to interpret Gibson's statement. Alternatively, Gibson might mean that the process (or the outcome) of affordance perception relates properties of objects to bodily properties of subjects, as in the width of an object being related to the subject's hand span and the resulting affordance perception is: *this is grasp-able*. So, either the body determines the affordance perception, or the body is one relatum in the relation whose outcome is the affordance (for a discussion of other interpretations of Gibson's theory of affordances, see ch. 3.5). The problem with Gibson's idea of affordances as external but involving a subjective aspect is that the distinction does not make sense anymore if you want to integrate both aspects in one kind of property. But this is exactly what Gibson tries to do, and is unfortunately not very specific about the fine-grained structure of affordances understood in this way.

It becomes even more complicated with Gibson claiming that

> the affordance of something does *not change* as the need of the observer changes. The observer may or may not perceive or attend to the affordance, according to his needs, but the affordance, being invariant, is always there to be perceived. An affordance is not bestowed upon an object by a need of an observer and his act of perceiving it. The object offers what it does because it is what it is. To be sure, we define *what it is* in terms of ecological physics instead of physical physics, and it therefore possesses meaning and value to begin with. (Gibson 1986, 139)

This refers to both aspects of affordances of the environment being *objective* and being *directly perceivable*, as in the process of affordance perception does not add meaning by means of mental operations to neutrally perceived objects. In this formulation however, the whole idea of affordances sounds a lot "more objective" and external than in other passages – it is hard to imagine where there is room left for integrating a substantial subjective aspect that is essential for something being an affordance. To sum up, affordances are objective in that they are determined by the properties of objects, but they should also be subjective to the extent that the bodily constitution of the perceiver should play an essential role, though it remains relatively unclear what exactly this role might consist of.

The other characteristic of affordances is their direct perceivability. As shown in chapter 3.2, the direct perceivability is connected with Gibson's definition of affordances and his theory of information pickup, however some important aspects of direct have not yet been mentioned.

The idea of direct perception is an expression of Gibson's "anti-cognitivism". He rejects the idea that perceiving is a mental process, involving computation, representation or image-like entities that are associated with sense-data of any description. He also strongly rejects the ideas of behaviorism, which explains animal behavior solely in terms of conditioned stimulus-response relations. Explaining behavior of animals in terms of reaction to stimuli appeared arbitrary to Gibson, regarding what to count as stimulus and what to exclude in the behavioral explanation. This is of even more relevance for the application of behavioristic explanation outside the controlled conditions of the lab, which lead Gibson to conclude that this notion of stimulus becomes too broad and meaningless. Also, the concept of a stimulus, which is central to behaviorism, implies that stimuli are discrete events/entities that have a determinate temporal extension. The problem becomes evident for Gibson in the case of perceiving persistent objects, as the experience of permanence cannot be stimulus mediated: if the sensory system is exposed to a permanent stimulus, the response of the receptor decreases and sensory adaption happens. Thus, the perception of object persistence has to be caused by something other than a stimulus that cannot be permanent without rendering itself unperceivable (cf. Gibson 1986, 56).

Mentalism – the view that perception and action are based on or caused by underlying mental states – is problematic because the observer does not perceives the external environment, but ultimately perceives (the content of) mental states and thus being detached, "cut-off" from the environment. This is unbearable for Gibson and needs to be overcome by a more adequate theory of perception. He would even go as far as to prefer behavioristic explanations of behavior over mentalistic, although assumed deficient:

> The doctrine of stimuli and responses seems to me false, but I do not on that account reject behaviorism. Its influence is on the wane, no doubt, but a regression to mentalism would be worse. Why must we seek explanation in *either* Body or Mind (sic!)? It is a false dichotomy. (Gibson 1986, xiii)

Rejecting mentalism and behaviorism as viable explanatory schemas, Gibson has to provide an alternative explanation of how perception and processing of action-relevant features of the environment works. He suggests that observers simply "read off" the information useful and necessary for interaction with the environment. This "reading off" can only work because the information about interactional features is objectively available – implying that affordances are real properties in the environment and *not* the result of perceiving neutral properties (form, color, shape etc.) and mentally inferring a function on that basis. Affordances are directly specified in stimulus information, everyone can in principle perceive them as they are really a part of the physical environment:

> The perceiving of an affordance is not a process of perceiving a value-free physical object to which meaning is somehow added in a way that no one has been able to agree upon; it is a process of perceiving a value-rich ecological object. Any substance, any surface, any layout has some affordance for benefit or injury to someone. Physics may be value-free, but ecology is not. (Gibson 1986, 140)

Gibson suggests not discussing the ontological status of affordances, as they are defined as objectively existing for him anyway, but rather focus on the question "whether information is available in ambient light for perceiving them." (Gibson 1986, 140) How can information for affordances be available in the light, especially more complex affordances such as "being good to eat"? It is one thing to assume that very basic information is given in terms of simple reflectance structures, such as a strong dark/bright contrast specifies an edge or a cliff – which is something to avoid running into or falling down. But most everyday affordances seem to be much more specific – how does the affordance of the door knob or door handle become available in the light, specifically as there are many ways to open doors. Gibson assumes that complex affordances are just compounds of invariants, forming a new single invariant – to avoid the need for a mind mentally combining the perceived individual invariants. Hence, even highly complex affordances are specified by invariants in the structure of the ambient energy flux, which renders them, in principle, directly perceivable as simple affordances. Gibson is certain though about the direct perception of *basic affordances* (cf. Gibson 1986, 143).

Another aspect in Gibson theory of direct perception is his view on the connection between learning and perceiving. Gibson claims that perception is an ability that is learned and can be more and more refined throughout ontogenetic development.

> The inputs of a special sense constitute a repertory of innate sensations, whereas the achievements of a perceptual system are susceptible to maturation and learning. Sensations of one modality can be combined with those of another in accordance with the laws of association; they can be organized or fused or supplemented or selected, but no new sensations can be learned. The information that is picked up, on the other hand, becomes more and more subtle, elaborate, and precise with practice. One can keep on learning to perceive as long as life goes on. (Gibson 1986, 245)

The process of learning mentioned here consists of being able to pick up different information. It is not the ability to make more sense of what one perceives, but to perceive more things, more different and complex information that has different and new meanings. Traditional accounts would stress the importance of establishing new connections between what one has perceived in the past and what is being perceived in the present and will be perceived in the future, enriching memory and allowing for new connections to be formed. This process of associating memory with sensory input is explicitly rejected by Gibson, claiming that the core fallacy of this view is that there is no explanation for why a certain sensory input becomes associated with a stored perceptual representation. There has to be a rule or a mechanism that would determine which kind of sensory input could be associated with which kind of memorized perceptual inputs – a new sensory input about a tree has to be associated with stored representations of trees in order to retrieve differences and similarities. Thus, the reason why the appropriate associations are established according to the traditional view of perceptual learning is because the sensory input is already categorized and then associated with a memory of the right category. This reasoning is circular for Gibson, stating that all "forms of cognitive processing imply cognition so as to account for cognition" (Gibson 1986, 253). In this example, the process of learning is explained on the basis of already existing knowledge – acquiring knowledge presupposes existing knowledge, and this in turn can either be innate or acquired. These problems should be overcome by taking the alternative route to perceptual learning which Gibson proposes. Perception, it is claimed, is an ability to pick up (objectively existing) information; learning consists simply in improving the ability to perceive in order to pick up increasingly subtle and complex information:

> Perception may or may not occur in the presence of information. Perceptual awareness, unlike sensory awareness, does not have any discoverable stimulus threshold. It depends on the age of the perceiver, how well he has learned to perceive, and how strongly he is motivated to perceive. (Gibson 1986, 57)

Important for Gibson's view that his theory of information pickup is not assigning any central role to memory at all:

> Evidently the theory of information pickup does not need memory. It does not have to have as a basic postulate the effect of past experience on present experience by way of memory. It needs to explain learning, that is, the improvement of perceiving with practice and the education of attention, but not by an appeal to the catch-all of past experience or to the muddle of memory. The state of a perceptual system is altered when it is attuned to information of a certain sort. The system has become sensitized. Differences are noticed that were previously not noticed. Features become distinctive that were formerly vague. But this altered state need not be thought of as depending on a memory, an image, an engram, or a trace. An image of the past, if experienced at all, would be only an incidental symptom of the altered state. (Gibson 1986, 254)

It is not entirely clear why, and on what grounds the perceptual system is able to attend to these new sorts of information and notice differences and new distinctive features *without* (mentally) comparing them with previously perceived information. A possible way of accounting for this process of attuning and sensitizing could probably be his concept of informational externalism. Information conceived this way is totally independent of the observer and specifies real properties. If the information is presented in an appropriate way, Gibson assumes that it can be adequately picked up by the observer. If new aspects are available in the right circumstances, the observer might thus be able to pick up these new aspects. This interpretation of Gibson's take on perceptual learning though is more an educated guess than a proper description of how this process might work in detail; unfortunately Gibson does not provide a more fine-grained account of this concept.

What does this imply for the directness of affordance perception? On the one hand, (basic) affordances should be perceived directly, on the other hand, perceiving affordances is an ability that is subject to the process of learning and development. Moreover, Gibson claims that perceiving affordances of objects comes first in perceptual development, children learn at later stages of their development to discriminate other properties of the object, such as surface, color and form. Children first discover the meaning of objects, therefore, the other perceivable aspects that are not affordances, are acquired through learning. In another passage, Gibson states however, that if

> the affordances of a thing are perceived correctly, we say that it looks like what it is. But we must, of course, learn to see what things really are for example, that the innocent-looking leaf is really a nettle or that the helpful-sounding politician is really a demagogue. And this can be very difficult. (Gibson 1986, 143)

From this it could follow that affordance perception is also learned and therefore not as direct and immediate as Gibson wants it to be. It seems that affordances could be the result of a process of finding out what an object actually is, and it is quite difficult not to conceive of this process as being cognitive. Gibson does not make clear what the difference between being able to perceive an affordance and learning how to perceive an affordance is, in case there is one. The mere possibility of having to learn what a thing really is, and therefore discovering what it truly affords (the innocent leaf affords something different than a nettle) seems prima facie to be a cognitive process and not one of simple pick up – as there is quite likely more to find out about the perceived thing, which might not be directly given in its appearance.

The passage just quoted leads directly to the problem of misrepresentation or misperception, the latter term being preferred by Gibson. The fact that perception in general and affordances in particular are subject to learning matters for Gibson's treatment of misperception. Every theory of perception has to account for misperception, in that it explains why we sometimes perceive things as different from what they really are. A cow in the twilight can be easily mistaken for a horse, a stick may appear bent in the water though it is perfectly straight, and looking at a white surface after staring in bright light may appear reddish to the observer. If perception is direct in that it consists of the direct, unmediated pickup of objective properties of the environment, errors in perception are rather unlikely to occur. Being aware of that caveat, Gibson proposes to think in terms of misinformation rather than in terms of misperception. This is a shift from the subjective failure to correctly represent or interpret the available information, as traditional accounts of perception would have it, to objective facts or an external cause: the information not being specific enough or ambiguous.

According to this view, "if information is picked up, perception occurs, if misinformation is picked up misperception occurs." (Gibson 1986, 142) The act of perceiving is always of the same type in both cases, what is variable and hence different is the information available.

3.4 Problems in Gibson's Account of Affordance Perception

Gibson's ecological account of affordance perception provoked a lot of critical reactions, as well as being well-received by the ecological psychology community, which made (and still makes) an effort to defend Gibson's claims against critics. Gibson wanted to overcome the cognitive science and psychology of his time and establish a radically new way of thinking about perception and conducting psychological research in general. Such an enterprise is naturally bound to polarize. Gibson inspired many reactions of rejection and severe criticism, as his claims were not only strong, but also sketchy in nature. Unfortunately, Gibson died shortly after publishing "The Ecological Approach to Visual Perception" and was not able to respond to his critics anymore nor able to provide a more detailed account of the controversial passages – this job has been taken over by his fellow peers, whose contributions will be reviewed in chapter 3.5.

This chapter will discuss the main lines of criticism found in the literature and in addition provide a few further arguments why the notion of affordances as construed by Gibson is implausible and needs revision in order to be a substantial philosophical and psychological concept. The central and most influential counter-arguments to Gibson's affordances have already been provided in Fodor and Pylyshyn's (1981) paper, which I will take as a starting point to move on to other more recent lines of criticism, yielding at a general discussion of Gibson's problematic account of perception and his account of affordances. This will provide the basis for discussing the accounts that were developed post-Gibson, mainly to overcome problems and elaborate the system Gibson had just begun to develop. We will see more clearly which of the other accounts, in their attempts to provide a substantial and extended notion of affordances, are able to handle the problems or are problematic in the same or other respects. Understanding the problematic elements in Gibson's theory will also help to develop an account of affordances that preserves its originally desired explanatory value but avoids common pitfalls and stands on more reliable theoretical foundations – thus rendering the concept of affordances more scientific and thereby applicable in all sorts of scientific behavioral research.

To start with, this is what I take to be the essence of Gibson's concept of affordances, basically the sub-claims contained in the two major claims, that affordances are objective properties and they can be directly picked up:

- Affordances are objective, they are properties of the environment
- Affordances are properties with reference to the observer.
- Reference to the observer consists of referring to the observer's bodily constitution, such as grip-size, leg-length etc.
- They are not properties of the experience of the observer or in any other way "added" to neutral properties of the environment.

- Perception is direct: it consist of a perceptual system picking up objective information, no other cognitive or mental states are involved in perception.
- Affordances are real properties and are (therefore) directly perceivable.
- Misperceiving affordances entails the pick-up of misinformation.
- Perception is an ability that is subject to development by learning to pick up more complex and subtle information.

Fodor and Pylyshyn (1981) discuss Gibson's account of perception in the light of what they call the "establishment view" and which can be described as an information theoretical view that claims "that perception [...] depends upon *inferences*" (Fodor and Pylyshyn 1981, 139). I will focus on one aspect of their criticism of Gibson's account, which is basically a variation of the 'poverty of the stimulus' argument: no visual input stimulus can be the bearer of all the information that the actual percept, or perceptual mental state contains.[10] It has been one of the standard arguments against behavioristic theories brought forth prominently by Chomsky (1959) in the second half of the 20th century and proven to make a strong case against all stimulus-response based theories of behavior. Gibson is anxious to dissociate himself and ecological psychology in general from mere behavioristic approaches, but, with Gibson claiming that all the information necessary and available for perception is already contained in the ambient light array, Fodor and Pylyshyn hold that the same line of criticism will be a strong case against the ecological framework. Fodor and Pylyshyn's criticism is mainly directed against Gibson's claim that perception is unmediated by mental processes. In their view, Gibson line of thought is in line with the behaviorists' claims that mental processes do not play a role, –let alone a significant one, – in explaining behavior, as the whole explanation can be given in terms of stimulus response and

10 An explicit reference to the „poverty of the stimulus" argument can be found here: „The consequence [...] is that visual perception typically involves inference from the properties of the environment that are (to use Gibson's term) 'specified' by the samples of the light one has actually encountered to those properties that would be specified by a more extensive sample. This sort of inference is required *because the causally effective stimulus for perception very often underdetermines what is seen*" (Fodor and Pylyshyn 1981, 142; my italics)

conditioned reflexes.[11] After discussing their attack on Gibson's account of direct perception, a discussion of more general criticism on the concept of affordances will follow.

One of Fodor and Pylyshyn's main arguments against direct perception in Gibson's sense (and hence against the direct perception of affordances) is that Gibson fails to provide an account of the information, i.e., the ecological properties that are directly perceived or "picked up". The only properties of the environment that could be possibly directly perceived would have to be "projectible", a kind of "property in virtue of which things enter lawful relations" (Fodor and Pylyshyn 1981, 146). Ecological properties, thus, have to be projectible properties, which mean properties that can be expressed by predicates that appear in laws. As example serves the common generalization: "all mammal have a heart" in contrast to "all mammals are born before 2016", where the latter predicate certainly is true *now,* but fails to establish a lawlike relation as it is true by coincidence, whereas the former generalization gives us good reasons to believe this is true in general (because having a heart is a defining thus lawlike criterion for being a mammal). The projectible ecological properties Gibson needs then "would be the ones which are connected, in a lawful way, with properties of the ambient light […] [these being] the projectible properties, and only those, that are the possible objects of direct visual perception" (Fodor and Pylyshyn 1981, 147) And exactly this is what Gibson fails to provide, considering his construal of affordances as paradigm ecological properties for direct pick-up:

> There are, for example, presumably no laws about the ways that light is structured by the class of things that can be eaten, or by the class of writing implements, though being edible or being a writing implement are just the sorts of properties that Gibson talks of objects as affording. The best one can do in this area is to say that things which share their affordances often […] have a characteristic

11 "The problem that we are raising against Gibson is, to all intents and purposes, identical to one that Chomsky (1959) raised against Skinner. […] Chomsky's critique thus comes down to the correct observation that there is no reason to believe that anything physically specifiable *could* play the functional role vis à vis the causation of behavior that Skinner wants controlling stimuli to play; the point being that behavior is in fact the joint effect of impinging stimuli *together with the organism's mental states and processes.*" (Fodor and Pylyshyn 1981, 143; footnote 2)

> shape (color, texture, size, etc.) and that there are laws which connect *the shape* (etc.) with properties of the light that the object reflects. But, of course, this consideration does Gibson no good, since it is supposed to be the affordances of objects, not just their shapes that are directly perceived. In particular, Gibson is explicit in denying that the perception of the affordances of objects is mediated by inference from prior detection of their shape, color, texture, or other such "qualities. (Fodor and Pylyshyn 1981, 147f)

From here, it follows that if there are no projectible properties that enable subjects to directly perceive that something affords to be eaten, to sit on, to write with etc., it is hard to explain why these properties should be perceived in a direct way at all; and not as, the inferential account would have it, be inferred from the basic visual properties perceived in accordance with, e.g. stored representations in memory, or any kind of similar explanation. More generally, this line of criticism can be understood in the same way as the criticism brought forth against behavioristic accounts: Either a stimulus can be lawfully related to a type of behavioral response, or the claims made by mid-20th century behaviorists are trivial at best, or invalid. If there is the possibility to respond differently to a stimulus of a certain type, then the lawful relation of stimulus and response is broken and the explanatory value is lost. The same for affordances: to be directly perceivable and thus to play a significant role in psychological explanations and theories, the lawful connection of stimulus (the information specified by the ambient light array) and affordance pick-up must hold. But as soon as there is a possibility to see one and the same object as affording something different, the information specified in the light cannot be lawfully connected to the affordance it ought to specify anymore. If the property of being edible is not lawfully related to visual properties of, say, an apple, then how can any affordance be lawfully specified by visual properties at all? The important part here is not that we could not detect or perceive affordances at all, but that the affordance perception cannot be direct and thus has to be explained otherwise – the inferential account of perception being only one possibility here.

This argument against the direct perceivability of affordances can be enriched by Fodor and Pylyshyn's claim that Gibson, seeking to establish an account of perception devoid of mental representations, cannot account for intentionality at all. Perception, as Gibson understands it, is merely extensional; cognitive phenomena such as belief, desire etc. are intentional and thus have to be accounted for in any psychological theory. Their argument in more detail is as follows:

> that (a) the prototypical perceptual relations (seeing, hearing, tasting, etc.) are extensional (and even where they are not, Gibson, in effect, treats them as though they were); (b) whereas, on the contrary, most other prototypical cognitive relations (believing, expecting, thinking about, seeing as, etc.) are intentional; and (c) the main work that the mental representations construct does in cognitive theory is to provide a basis for explaining the intentionality of cognitive relations. (Fodor and Pylyshyn 1981, 188)

This argument is only a problem for Gibson, if one can show that affordance perception cannot be explained on the basis of purely extensional relations. As Fodor and Pylyshyn claim, only *seeing* can be explicated in terms of an extensional relation, but *seeing as* is always an intentional relation. Although some core aspects of perception might be explicated in terms of extensional relations, some other aspects can only be understood in intentional terms. Based on that distinction, it is hard to see how an account of affordance perception can be given in non-intentional terms. Perceiving the edibility of an apple means seeing the apple as edible – in contrast to perceiving its paperweight-affordance or its throw-ability or seeing the apple as something that keeps the doctor away. As a consequence, the only option for Gibson to explain the apparent many options of what an ecological object can be *seen as* is by ascribing them many different properties, instead of going down the representational road and describing one and the same property of the object as being represented in different ways. This alternative way of conceiving of different aspects of one's environment would not be problematic if Gibson would offer a convincing account of how these properties (the different affordances of the objects) are specified in the ambient light array, without presupposing that only what is picked up is an affordance and therefore being picked up directly per definition.[12] How is it, that an apple's property of being edible can be specified in the ambient light array differently than the property of being graspable or being throw-able? As Fodor and Pylyshyn write, "*Property* is an *intentional* notion in the sense that coextensive sets may correspond to distinct properties [...] however, *specification* is an *extensional* notion" (Fodor and Pylyshyn 1981, 191). Seeing the apple as edible implies edibility being specified in the light, if the apple is also throw-able, then edibility and throw-ability will be specified in the same way in the light – to put it differently, there is a property x of the ambient light array that specifies both edibility and throw-ability (not to mention the property of keeping the doctor away). It is hard to tell with such an account of perception when and how different properties are picked up, when they are specified in the light in the same way. Accounts of perception and cognition that assume mental representations as central elements of cognition on the other hand are able to deal with these kinds of differences by appealing to different ways of representing the same property. If this analysis is valid, Gibson's account faces serious difficulties, as he is unable to account for the phenomenon of intentionality in perception and fails to provide a convincing explanation for perceiving different aspects of the same object.

Consequently, it follows that affordances cannot be directly picked up – at least not in any remarkable sense. Maybe some simple, very basic affordances could be picked up in a way that could justify the description 'direct'. What could qualify for this sense of directness could be simple affordances that have a strong correspondence relation to bodily features – the height of a doorframe, that determines walk-through-ability, or a

12 Furthermore, Fodor and Pylyshyn are right in pointing to the fact that Gibson *can* only claim that what is directly picked up is light as such – light being the only possible kind of stimulus the sensory system could resonate or attune to. This of course provokes the further question how to explain the pickup of affordances without inferential processes if only structured light can be picked up. The only way out for Gibson is claiming that affordances nomologically or reliable covary with structure in the light and thus the pickup of structured light *is* the pickup of the respective affordance, which, once again leads to the question how one and the same object can structure the light differently so that different affordances can be picked up. (cf. Fodor and Pylyshyn 1981, 159ff)

handle that is in a hand related size and orientation. These simple affordances might elicit a very immediate, almost automatic motor-response and could thus count as 'direct'. This directness though is no longer valid for affordances of objects that are not correlate-able with primitive movements and bodily features, which is the majority of affordances and thus the interesting cases. In chapters 6 and 8 of this book, there will be more on the distinction of simple action-related representations to more complex one and how directly they are correlated or respond to bodily features and simple movements.

If affordances cannot be fully specified in the ambient light array and thus not be directly picked up, does this have any impact on Gibson's ontological specification of affordances as special properties that cut through the objective/subjective dichotomy (by being either both or neither objective or subjective)? A major difficulty to giving a satisfying answer to this question lies in the rather vague and ambiguous way Gibson talks about affordances – in fact his claims seem controversial if not inconsistent. On the one hand, Gibson is eager to stress that affordances, being values and meanings are objective properties of the environment which just have to be picked up and thus exist at least partly in an observer-independent way. On the other hand, Gibson claims that affordances are properties of the environment that refer to properties of an animal, which gives rise to a dispositional and/or relational interpretation of affordances. The reference to the animal should take place in terms or referring to physical features of the animal, such as leg-length, which determines e.g. the possible steps an animal can climb. It seems as if Gibson was indecisive as to whether he should consequently follow his direct realism or shift the focus towards more constructivist conception of affordances: the idea that affordances are subjective in that they are determined by the individual subject and therefore only existent in the subjective world. There is a certain tension in Gibson's account that cannot easily be resolved – interpreting it either way leads to serious difficulties.

What complicates matters further is Gibson's premise (and goal) to exclude all kinds of mental processes in the detection of affordances (and his psychology of perception in general), which is difficult to reconcile with his claims about learning to perceive and learning new affordances. There is a certain tension in this way of defining affordances, and this is mostly due to Gibson not being specific about contextual features and especially the role intentions could play in the bigger picture. The notion of affordances as defined by Gibson, lacks an intentional, motivational element which is able to explain why an animal acts on the presence of certain (objective) properties in the environment. It is not entirely clear why Gibson avoids considering intentions in his theory – most likely it is because generally intentions are held to be a type of mental state or representation and therefore in conflict with his insistence that perceiving affordances should not be seen as based on mental processes. The need for an intentional element stems mainly from the reason that every object has, in principle, infinitely many properties which can give rise to action possibilities for animals, and every animal has many ways of interacting with the environment. This implies that an object can e.g. be grasped in many different ways, as it possibly features more than one "handle", and animals can have different ways of grasping as well. Moreover, every object can afford different actions in different situations, being used for different or entirely new purposes. Affordances can even be invented, such that animals can come up with or find out about new ways of using already known objects. In recent studies New Caledonian Crows and Keas show flexible problem solving behavior and tool use: they are e.g. shaping a hook by bending a stick in order to reach for food (cf. Weir et al. 2002; Weir and Kacelnik 2006; Auersperg et al. 2011). It is thus mainly the need to explain flexible behavior that demands for a further element, one which neither classical stimulus-response behaviorism nor Gibson's theory of direct information pick-up can deliver. For Gibson, what specifies affordance perception is given in the animal's context and the current situation; intentions are merely belonging to the contextual features and need not be further paid attention to or to be analyzed separately.

However, this is a problematic view, as intentions cannot (and should not) easily be subsumed under general objective contextual environmental features – the concept of intention would thus become superfluous and the explanation circular: if intentions are part of the environmental and contextual properties, then every action can merely be explained post hoc by claiming that the animal just has had the corresponding intention, otherwise it would not have acted on the affordances it actually acted on but would have acted differently. If I want to know why someone turned the light on, it would not be very satisfactory if the answer would be: because he intended to. So an action explanation cannot merely state that an intention for action must have existed, because otherwise the action would not have been executed, or the explanation would be circular. An intentional explanation adds something significant, such as the information that it was too dark to read or that the person wanted to check if the light bulb is still working. An intention (among other factors) can rationalize, i.e., give reasons for actions (cf. Davidson 1967), but these reasons would be meaningless if they would just state the obvious, and this would be the consequence of subsuming possible intentions in the general context, to avoid having mental representations in the theory. It is exactly what we want an explanation for: why do subjects act on specific, different affordances in different situations, especially when their behaviors show a high degree of flexibility. Intentions can be an explanation for this: they guide the perception of affordances, by selectively attending to those affordances that match the intended goal states.

Without considering intentional states at all, Gibson cannot explain why animals sometimes act upon some, and not on other affordances. Given that there are infinitely many possible affordance and that, according to Gibson, they are already specified directly in the ambient energy flux, there must be an additional reason why the animal picks out some affordances and not others. Gibson could of course deny that behavior is flexible in some species and therefore the problem of affordance selection is no real problem. However, this would destroy the explanatory benefits of affordances and the whole account would consequently collapse into a simple stimulus-response behaviorism of action explanation.

The bottom line of this reasoning is that without adding an intentional, motivational element to the theory of affordances, the desired explanatory value is corrupted. There must be some sort of mechanism for all animals that guides attention to affordances. Imagine a squirrel with a nut sitting on the branch of a tree. In its vicinity, there is a 'jump-to-next-tree' affordance, a 'climb-up-tree' affordance and a 'hide-nut' affordance. Whatever it is going to do depends partly on the action possibilities the squirrel will perceive, and there are plenty of them. Some guiding mechanism has to guide the relevant affordance perception, otherwise it remains entirely unclear why the animal acts on certain affordances and not on others. One might be reluctant to attribute full-blown intentions to squirrels, cats and toddlers, but speaking of motivations that influence and guide the affordance selection should be less controversial. However, if this is a convincing argument, then it is difficult to see how affordances could be characterized as objective, when affordance perception is mainly driven by intentional states.

Gibson thought there would be a way out of this dilemma by defining affordances as neither an objective property nor a subjective property or being both facts of the environment and facts of behavior (cf. Gibson 1986, 129). The subjective element according to this definition would be the body of the animal – an affordance is a property of the environment relative to the physical constitution of the animal. The physical constitution plays an important role in what an animal could consider as an action possibility (for itself), by relating physical properties of the environment to physical properties of the animal. But describing the animal-environment relation this way leaves us with a mere relation between objective physical properties, while a proper subjective element remains entirely absent in this relation. Thus, affordances defined in terms of a relation of mere physical properties are not able to explain behavior anymore. The fact that a subject is of a certain height or weight, does not explain why the subject acts upon certain environmental properties. Just because someone can lift a heavy box does not mean that the person is actually going to lift the heavy box. For explaining behavior, something else must be added, such as a proper subjective element. Defining bodily features as subjective is not enough, in addition there must be a proper subjective element establishing or initiating the affordance relation. Otherwise, the explanations would empty, as in "Why did you reach for the bottle? Because I can." "Why did the cat climb up the tree? Because its physical properties related to the tree's properties establish climb-ability." It is obvious that this is not what one expects from an action explanation. Subjects can do all sorts of things merely by their bodily constitution, but this does not entail that they will do all of this, so something like a basal motivation, an intention or an action-goal representation has to be added to the affordance relation to become explanatory significant.

Accordingly, it follows that both Gibson's statement that affordances are physical properties and the claim that meanings and values are external to the perceiver are rather incomprehensible. It is one thing to claim that some physical property can be more or less directly specified in the light array, but it is a much stronger claim that the value of objects in an animal's environment is external in that way too. If an object is of some value for an animal or has some meaning for the observer, then this is because the object matches the desired goals and purposes of the observer. That means that the object's properties, constituting partly the affordances for the subject, can only be one factor in the whole complex that comprises the value or the meaning for the subject.

Are there objects that can have an objective value? Presumably, one could speak of an item of food, say some sort of nut, has a meaning for a squirrel which is not based on the squirrel's intention or other 'mental' states. The nut is food, it has nutritional value and whenever a squirrel encounters a nut, it will try to take it or eat it. The nut has this objective value for the squirrel only because the squirrel has an inbuilt nut-detector that triggers a certain behavioral pattern every time the right perceptual input is processed by the nut-detecting system. In this sense, the squirrel is determined to react to nuts with the same behavioral pattern over and over again, this being not an example of flexible behavior anymore. Still, the value of the nut is not entirely external to the squirrel, as it is the existence of the nut-detector that makes the nut valuable. The nut has no value for animals unable to digest nuts and thus lacking any nut detecting systems. The idea that meanings and values could be external to the perceiver becomes even more problematic in more complex actions. Imagine someone camping in the wild, intending to secure the tent with tent pegs but forgot to bring a hammer. A short look around should be sufficient to find a substitute, say a stone or a big piece of wood. The stone affords hammering the tent peg into the ground too. But can its value be external to the camper, if he was the one looking for an object that fulfills a certain purpose, meaning that he "knew" which properties he was searching for? This is actually difficult to conceive and it seems to be much less problematic to assume that an object can only have meaning in relation to a general purpose or goal, which can only be an intentional state of an agent. The squirrel is determined by his instincts, based on a genetically determined, evolutionary selected mechanism to collect nut-like objects. But this is most likely a hardwired behavior routine which does not allow for much behavioral flexibility; therefore the explanation will be rather simple. It could possibly be given in a functional description such as: whenever animal of type squirrel (in a state of being hungry) encounters nut-like object (specified by visual features) it will do (x,y,z). Therefore, a purely behavioral description in a stimulus-response style will be already quite complete; at least the need for introducing a new kind of property which entails problematic ontological commitments seems not to be necessary at this stage. Explaining complex actions on the other hand with affordances as physical, external properties does not yield a satisfactory explanation either, because the proper subjective part that explains the property selection by the agent is excluded by Gibson's concept of affordances and is therefore deficient.

3.5 Gibson's Successors

Many attempts have been made since Gibson's death to interpret, revise and save the concept of affordances. This section will provide a brief discussion of the most important accounts, critically evaluating if the revised versions can overcome Gibson's problems (as analyzed in the previous section) and thus make the concept of affordances scientifically applicable. The accounts that will be discussed can be roughly divided in two views: Affordances as dispositions, and affordances as relations. Thus, the dispositional view postulates that affordances are nothing but intrinsic properties of the environment and the animal, whereas in the relational view, affordances are something more in the sense that they are synergetic or emergent properties that arise out of the animal-environment relation. The dispositional view of affordances is held by Turvey (1992) and

Scarantino (2003), whereas Stoffregen (2003), Heft (1989) and Chemero (2003) defend the non-reductive version. Siegel's account does not really fit into either category, so it is treated as a standalone contribution to save the notion of affordances, as well as the short section on Norman's notion of perceived affordances. Another account that will be discussed in this section is Nanay's (2011) concept of 'action-oriented perception'. Although not explicitly referring to Gibsonian affordances, Nanay's account shows some striking similarities to the original affordance concept.

3.5.1 Affordances as Dispositions

3.5.1.1 Affordances as real possibilities

As mentioned above, Gibson's definition of the properties involved in the animal-environment relation is rather vague. One possible way of clarifying the notion of affordances is to understand them in terms of dispositional properties. In this regard, Turvey conceives of an affordance as being "a particular kind of disposition, one whose complement is a dispositional property of an organism" (Turvey 1992, 179). Affordances are real possibilities, in that they, understood as possibilities for actions, "constitute an ontological category, not an epistemological category" (Turvey 1992, 174). The affordance is the disposition that needs to be complemented by what Turvey calls 'effectivity', which is a dispositional property of an animal. Interchangeable as it is, the affordance could well be a disposition of an animal to behave in a certain way that needs to be complemented by dispositional properties of the environment. Crucial to understanding Turvey's interpretation of affordances is his notion of dispositions or being a dispositional property. A disposition is defined as being "tantamount to an actual state of affairs minus particular conditions" (Turvey 1992, 179) that will become actualized when certain conditions are fulfilled or present. To have the disposition of being water-soluble in this respect means something has the property to dissolve when getting in touch with water. Water is the condition that provides actuality for the disposition – though the thing in question still has the dispositional property of being water-soluble when there is no water present. Therefore, dispositional properties cannot exist independently of facts or features potentially provided by the environment: "Complementarity occurs in the very definition of a dispositional property." (Turvey 1992, 178) Having a certain disposition always entails the conditions in which a certain state of affairs will be actualized.

3.5.1.2 Dispositional predicate analysis

Scarantino also proposes a dispositional analysis of affordances, stating first that the existing dispositional accounts fail to specify what kinds of disposition affordances should be (cf. Scarantino 2003, note 9). To overcome this deficit, he offers a semantic analysis of the predicates used to express the dispositional properties. Thus he states "that to clarify the meaning of properties is to clarify the semantics of the predicates (if any) expressing them" (Scarantino 2003). Scarantino adopts Mumford's notion of a dispositional predicate that depends on "the way in which its ascription entails subjunctive conditionals" (Scarantino 2003). In this regard, the ascription of some X being fragile can be formulated as "if X were (suitably) hit, then X would break" (Scarantino 2003) – the ascription entails the subjunctive conditional in a conceptually necessary way. Being fragile in this sense is what the subjunctive conditional expresses. Another important characteristic of dispositional predicates according to Scarantino is their incompleteness – they depend on some completing background circumstances; objects are inflammable, water-soluble and the like, given some background conditions. The specification of which conditions are relevant depends on which factors are taken into account, e.g. how broad or narrow the set of possible conditions is defined. Scarantino wants to exclude special cases of conditions under which e.g. steel can be soluble and considers solely what he calls "normal ecological circumstances" (Scarantino 2003). In identifying affordances with dispositional properties, the analogue holds that the predicates describing affordances, such as climb-able and reachable, etc., are dispositional predicates, so that they also entail a subjunctive conditional of the form: If at time t background condition C were the case, then a manifestation M involving X and O would be the case, where X is the affordance bearer and O an organism (cf. Scarantino 2003). Affordance predicates are "time-indexed incomplete predicates, whose completer is a set of background circumstances referring to an organism at a time in a set of environmental circumstances. For example, a tree X is climbable/not-climbable not simpliciter, but at time t relative to squirrel O in background circumstances C" (Scarantino 2003). The background circumstances could involve conditions such as the tree is not on fire or that the squirrel is physically free to move.

Scarantino proceeds in defining different types of affordances, namely goal affordances (*their manifestation is a doing*) and happening affordances (*their manifestation is a happening*). According to the doing/happening distinction in the philosophy of action, doings necessarily involve goal-orientated intentions, whereas happenings are events without intentions involved – things that just happen to an organism. As Scarantino takes intentions to be propositional, the doing/happening distinction can only sensibly be made relative to human organisms that provide the adequate conceptual organization necessary for having propositional intentions.

At this point, Scarantino departs from Gibson and proposes three kinds of affordances: 1. Basic physical affordances (a flying ball is catch-able), 2. Non-basic physical affordances (a flying ball is score-with-able), and 3. Mental affordances (a number is divide-by-two-able). The first type of affordance is the kind that can in principle be perceived by all organisms, including non-linguistic animals, whereas the latter two kinds are perceivable only by organisms with the right conceptual organization, which makes them language-dependent. Scarantino thus follows Gibson in stating that there exist objective, directly perceivable affordances, namely basic physical affordances, but expands the realm of affordances to more complex, higher-order affordances, whose perception is limited to higher-order cognitive organisms, involving conceptual knowledge and memory. He leaves open the question whether the latter kind of affordances is (directly) perceivable and if so, how they are perceived.

Neither Turvey (1992) nor Scarantino's (2003) take on affordances offers a satisfying solution to the major problems in Gibson's (1986) account. Treating affordances as mere dispositions, as Turvey does, does not explain why behavior is executed on some occasions, but not on others. The only way for Turvey is to put too much weight on the conditions needed to actualize a dispositional state. If behavior is accounted for in analogy to the water-soluble case, this would be a reductionist understanding of affordances, explaining behavior merely with the occurrence of certain states of affairs. An affordance would be reduced to the disposition of, e.g., being able to grasp bottles and the presence of a bottle of the right size. From this, it does not follow that the subject in this situation will actually reach for the bottle. There are always infinitely many possible affordances surrounding animals due to their being disposed to behave in certain ways, but this does not entail that the animal will act upon all these affordances. It is hard to see why this should yield a better explanation for (complex) behavior than any purely behavioristic description.

Scarantino's (2003) predicate analysis of dispositions is no real advancement either. The explanatory work is, once again, mainly done by the background conditions and what he calls "normal ecological circumstances". That a tree is climbable for a squirrel in a situation without conflicting conditions (e.g., the tree being on fire) is not enough to explain the squirrel's behavior. Only if one allows for motivational aspects to be part of the background conditions, the manifested behavior can be explained by also referring to the situation's affordances. In a room full of chairs, having the disposition to be able to sit on chairs of the given size, an explanation has to be given why a subject picks out one chair and not another. This can only be done by referring to subjective aspects, such as the subject's general preference to sit at aisles, to in the back rather than the front rows etc. Either, the background conditions will be overladen with all possible aspects in a situation and thus diminish the explanatory appeal of the dispositional analysis, or the specification of the conditions will always be vulnerable to leaving out relevant aspects.

Scarantino's distinction of affordance types on the other hand is an improvement: by recognizing that basic action affordances differ significantly from higher-order action affordances, Scarantino allows for different explications of the affordances on different levels. For instance, basic affordance perception can be explained by referring to cortical structures such as the two visual pathways, with the dorsal stream processing action-related object information (for a detailed discussion, see ch. 6). Higher-level affordances, as involved in pursuing more abstract, distant goals have to be explained with different cognitive mechanisms. Thus, Scarantino justifiably addresses a major flaw in Gibson's (1986) account: his general theory of direct affordance perception can never account for all types of behavior and action manifestation.

3.5.2 Affordances as Relations or Emergent Properties

3.5.2.1 Affordances and intentions

A second group of approaches treats affordances as relations or emergent properties of systems. The first one to explicitly conceive of Gibson's affordances as relations is Heft (1989). In treating affordances as relations, Heft tries to make sense of Gibson's claim that affordances are neither just objective nor subjective properties – they are synergetic properties that emerge from the animal/environment relation. According to Gibson, affordances cannot be objective properties in the strict sense, as they "are not specifiable independent of an individual, as are physicalistic properties such as mass and extension" (Heft 1989, 4). At the same time, although necessarily involving a perceiver, they are not purely subjective properties that reside in the mind of the perceiver, as they are conceived as ecological facts well in accordance with Gibson's anti-mentalistic framework. Given that, affordances don't belong to either of these two ontological categories alone, but have to be conceived of as being relational, which implies that their existence depends on the existence of both of the relata. As Heft states, the "hallmark of an entity with a relational quality is that its specification implies a second entity" (Heft 1989, 5). Affordances are these kinds of entities: "They are the environmental counterparts to the animal's behavioral potentialities" (Heft 1989, 6). Therefore, objects, which are smaller in size as the hand span are the other relatum (the "environmental counterpart") of the act of grasping, which is "only comprehensible in relation to a thing which may be grasped" (Heft 1989, 6). Entities with relational qualities, the affordances, "complete the unity of the behavioral act" (Heft 1989, 6) and specify goal directed action together with the related behavior. One of the key questions for an account of affordances that preserves their objective nature insofar as they are constituted by ecological facts (e.g. facts of the environment and the animal) is how to specify which affordances are going to be perceived by the animal and become relevant for behavior. With the facts of the environment always present, an animal has to single out only some aspects that matter for current behavioral possibilities – it is very unlikely, and would be very inefficient, to perceive all possible affordances at any given time. In addressing this question, Heft introduces the notion of intentionality (purposefulness): "Which particular affordances are utilized in a given environmental setting will depend on intentional processes of the perceiver" (Heft 1989, 10).

Affordances understood in Gibson's terms refer to certain bodily dimensions, e.g., steps afford stepping relative to the body scale of animals, in this case leg length, and doorways afford passing through relative to height and width of animals. Heft agrees with the
dependence on body-scales, but wants to go beyond a mere dependence on bodily features:

> However, I would like to suggest that the affordances of the environment refer to the body in a much more fundamental manner than mere body-scaling per se. Affordances are specifiable relative to what an animal can do, relative to what his potentialities tor action are. That is, the environment's affordances are to be defined in relation to the body as a means of expressing various goals or intentions. (Heft 1989, 11)

In conceiving of the body in a more phenomenological sense (following Merleau-Ponty), Heft wants to broaden the concept of affordances and introduce intentions and goals in addition to physical properties. Heft is sympathetic to Merleau-Ponty's notion of intentionality, and states:

> intentional acts are always situated. That is, inherent in an action is a reflection of a situation or a set of conditions. An intention is not describable in the absence of some foreseeable expression of it in the world. In this respect, intention does not refer to a mental representation; It is not a mentalistic notion. (Heft 1989, 11)

By adopting Merleau-Ponty's position, Heft attempts to enrich the concept of affordances with intentions and goals and at the same time preserves Gibson's anti-mentalism regarding the perception of affordances. Therefore, the combination of affordances, e.g., the "ecological resources for behavior" (Heft 1989, 12) and the physical properties of the animal define the scope of intentional acts that can be expressed. Moreover, an animal perceives affordances primarily in relation to its intentions or goals, and not only in relation to its physical properties. Hence, there are three crucial aspects for explaining the behavior of animals, which are interwoven and cannot be treated separately: the affordances of the environment, physical properties of the animal and the animal's intentions or goals. Which affordances are perceived is determined by the intentions of the animal, relative to its physical properties or action capabilities – different goals will make different affordances salient.

The problem with Heft's phenomenological interpretation is that his notion of non-mental intentions remains rather obscure. Acknowledging the problem of the missing proper subjective aspects in Gibson's affordance concept, Heft is committed to anti-representationalism and can therefore only introduce a non-mental notion of intention. Furthermore, Heft declares the non-mental intentions necessary for affordance perception, which is at odds with Heft's fundamental assumption of direct affordance perception. Moreover, it contradicts contemporary empirical evidence, showing that affordances are perceived and influence a subject's performance even if the affordances are task irrelevant, and therefore unlikely to be included in the subject's intentions (cf. Ellis & Tucker 2001; for further elaboration, see ch. 8)

3.5.2.2 Emergent properties of animal-environment systems

Another relational account to defining affordances is given by Stoffregen (2003). In his aim to propose a formal definition of affordances, Stoffregen initially rejects the formalization of the notion of affordances given by Turvey (1992, see above). Stoffregen argues that Turvey's account faces serious problems regarding the specification of affordances and direct perception. Any definition of affordances, he argues, has to be "compatible with a general theory of direct perception" (Stoffregen 2003, 122). Central to Stoffregen's account is the claim that affordances are only relevant in animal-environment systems. In a binary system, every component has certain properties, but the system regarded on the whole also has properties that may be distinct from the properties of the parts. These system properties are emergent properties because they are not properties of components of the system. Stoffregen illustrates that with the example of a triangle that is composed of three individual lines, where the lines and the triangle have distinct properties that cannot be reduced to one another: the triangles properties emerge from the properties of the three lines. In this sense, the animal-environment system has emergent properties that are not properties of either animal or environment. Affordances are defined as exclusive properties emerging from the whole system of the environment/animal relation. Although existing only in relation, affordances are ontologically "real" or objective properties that are "persistent, that exist prior to and independent of actual behavior" (Stoffregen 2003, 123). Stoffregen also introduces intentions for action to his affordance account, to address the problem why a subject acts on only some affordances out of a multitude of possible available affordances (cf. Stoffregen 2003, 125). In this interpretation, affordances exist independently from being actualized as well as intentions can exist without being satisfied or driving action. As the affordances should be emergent properties of animal-environment relations, the relations hold independently from being perceived or playing an active role in behavior:

> The persistence of affordances prior to their exploitation permits them to be specified and detected prospectively, which in turn permits affordances to function as the cornerstone of prospective control. (Stoffregen 2003, 126).

Their being independent, thus objective, enables affordances to be directly perceived and to function as set of actions a given intention will pick the appropriate action from. Intentions limit the possible affordances and vice versa: as behavior is the result of complementary affordances and intentions, not every existing affordance will satisfy a given intention, in the same way, a given set of affordances will give rise to some intentions only.

The notion of 'emergent properties' is far from being uncontroversial. A reductive materialist would definitely reject the idea of higher order properties that cannot explain by or reduced to basic level properties. But even for a moderate materialist, the notion of ontologically real properties that arise out of a system's structure and are not given already by the properties of the parts is at least peculiar. There is a less controversial understanding in terms of levels of description but this seems not to be what Stoffregen has in mind. Without going into detail, it can generally be said that it is possible to find examples for higher-order properties of a system that are not identical with its basic-level properties, while it is more difficult to demonstrate the ontological irreducibility of the different levels. Apart from the ontological difficulties any account of emergent properties faces, it is also unclear how conceiving of affordances as emergent properties of an animal-environment system provides a better understanding. With the premise that affordances exist prior to being perceived or acted upon, the nature of the relation in question becomes even more unclear. Stoffregen is thus committed to the claim that without necessarily perceiving it, a subject is always in relation to properties of the environment with affordances emerging form these relations. First of all, the question arises where to draw the boundaries of the animal-environment system. As perceptual contact is not necessary for the existence of the relation, a subject is in principle related to the whole environment. A defining criterion is needed, which restricts the possible relations, otherwise everything could be related to everything and the concept becomes meaningless in explanatory terms. Second, the idea that all possible affordances exist already and have merely to be detected and exploited is leading to strange consequences when considering more sophisticated behavioral possibilities. The brush and canvas in the room might give rise to the affordance of 'being-paint-with-able' in relation to subject able to grasp and hold the brush. It is less obvious to assume that the 'being-for-paining-a-truthful-copy-of-the-Mona-Lisa' affordance also exists as emergent property and can be detected or exploited, even with the complementary intention. It thus seems that Stoffregen's account is not able to add anything substantial to Gibson's (1986) account other than bringing in intentions for actions, which addresses one of the apparent neglects in Gibson's original proposal.

How Stoffregen can still maintain the notion of direct, unmediated pickup of affordances when intentions in his account drive affordance selection stays incomprehensible.

3.5.2.3 Affordance perception is feature placing

The final account of relational affordances is given by Chemero (2003), who's main aim is to give a definition of affordances "that makes them more ontologically respectable yet still does justice to Gibson's conception" (Chemero 2003, 182). Central to his criticism is the vague ontological designation of the properties identical to affordances, which are supposed to be neither objective nor subjective, but both (cf. Gibson 1986). Furthermore, Chemero also rejects most of the attempts to give formally and ontologically more adequate definitions (some of which have been discussed in this chapter). In particular, he rejects the dispositional analysis proposed by Turvey (1992) on the grounds that affordances are understood as properties, either of the environment or animals (cf. Chemero 2003, 183). Chemero's alternative definition of affordances understands affordances as relations of certain aspects of animals and of certain aspects of situations. The basic logical structure of affordances can be formalized like this:

> Affords-Φ (environment, organism), where Φ is a behavior. (Chemero 2003, 186)

Spelling this out, the two relata are the environment and the organism among which the relation 'affords-Φ' holds. This is analogous to other relations like 'Taller-than (Shaquille, Tony)' which means that Shaquille is taller than Tony. The relation holds only when both of the relata are present and is therefore dependent on the existence of the relata, which implies that neither of them inheres what the relation stands for (thus, 'taller-than' is inherent in neither Shaquille nor Tony). Although the affordances depend on the existence of the relata, and are in this sense not an "extra thing" in ontological respects, they nevertheless are real in the sense that they are perceivable, such as one can also perceive the fact that Shaquille is taller than Tony. To say that affordances are relations of environment and organism, one has to explicate which relata are related and how they are related. According to Chemero, environmental relata are features instead of properties, where the latter are predicated of objects, while features are ascribed to situations only (cf. Chemero 2003, 185). The other relata are an animal's abilities, which are functional properties of the animal's body. Affordance perception should be understood as *feature placing*, a notion from Strawson (1959), describing the recognition of certain situational features. Feature placing sentences, such as 'it's raining' or 'it's dinnertime', do not predicate a property of an object but rather state that a certain feature is present here and now (see also ch. 4 on causal indexicals). Accordingly, affordance perception is detection of an 'opportunity for do-ability x (for animal y with abilities z) now' (e.g. 'sitting opportunity (for animals able to sit) here, now'). Perception of affordances also has a relational structure, which looks like this:

> Perceives [animal, affordance-of-Φ]. (Chemero 2003, 191)

Normally, an animal just perceives the affordance, and not the relata the affordance involves. Information about abilities and about features are therefore not content of the perception, which reduces to perceiving what behavior is afforded for the animal. This can be illustrated by considering the phenomenology of affordance perception: normally, a subject would simply perceive stairs as step-on-able, without perceiving her stepping abilities or perceiving the riser height of the stairs. Affordance perception can thus be said to be transparent: all the subject perceives is the afforded action possibility and not the actual aspects instantiating action opportunity.

This is a valid description from a phenomenological perspective, but misses the point. Saying that the relation as such is not perceived is trivial, as subjects naturally perceive their environment in terms of higher-order features and not in terms of basic level properties. This is most certainly true for all perceptual relations: a perceptual relation (say, of perceptual system and an object of perception) also needs both relata to exist, as it would not make sense to speak of a perceptual state without being about any object. At the same time, what is perceived is simply the phenomenal object, and neither the object's surface texture, which is determining light absorption and emission, nor the properties of the perceptual system. In that sense, whatever is perceive is dependent on basic level structures, but the actual perceptual content is always of a higher level – the perceptual content is about cups, tables and chairs etc. This is not to say that Chemero is wrong, but that his description applies to all perceptual acts, without being committed to direct or inferential views of perception. Thus, this description will not secure the explanatory value of affordances, but neither will treating affordances as relations of features and abilities, at least not with regards to saving Gibson's (1986) general claims. This is because this relation can only explain simple behavior that involves a mapping of bodily features to environmental features. Examples are reachability, sit-on-ability, etc. Perceiving these simple action opportunities might be unmediated in the sense that memory and prior experience does not play a major role. Perceiving an object as reachable can primarily be explained as a function of the egocentric representation of action space, which is an automatic, subconscious process. For more complex examples, such as perceiving a plug of an electric device as plug-able into power points, subjects rely on stored knowledge as well as a property that is predicated to an object. Defining affordances as placed features severely limits the explanatory values of affordances, as most actions cannot be explained with affordance perception thus defined.

The advantage of Chemero's relational definition of affordances is that he is able to actually specify the rather nebulous Gibsonian subjective-objective definition of affordances by interpreting it as a relation that has to hold in order that the affordance be perceived. In this sense, Chemero can explain how affordances, as relations, can have a subjective and an objective aspect, with the subject having to be in the right kind of relation to environmental features to perceive an action opportunity at all. Unfortunately, his analysis does not explain how the subject, once a relation is established, extracts the relevant information provided by the relation that defines the affordance. It needs to be explained how the subject detects the action opportunity that is implied in the relation. A plausible way to explain it would be by inference on the basis of perceptually available information, which is related to previous experience stored in memory – however, this explanation is contrary to the direct perception claims made by Gibson and is therefore not viable for Chemero either. Thus, Chemero is only offering feature placing as a way of understanding affordance perception. However, feature placing as such is not able to account for direct affordance perception, as the directness of feature placing itself is questionable – to place a feature in time and space, one has to rely on already learned behavioral possibilities. To place a 'sitting opportunity', a subject has to have some prior experience with the action of sitting and also of objects that afford sitting, either first hand or by observation, which makes feature placing itself the explanandum rather than the explanans.

3.5.3 Non-Gibsonian Accounts of Affordances

In this section, I will discuss accounts that try to define affordances, but without explicitly referring to Gibson (1986) and also without the primary intention to improve or advance the concept of affordances in the light of Gibson's ecological premises. In that sense, these accounts do not share the premise that affordance perception has to be direct, but instead offer a representational interpretation of affordance perception, such as Siegel (2014). The aim is to show whether affordances outside of the Gibsonian framework can be defined in a way avoiding some of Gibson's problematic claims, while maintaining a special explanatory value.

3.5.3.1 Perceived affordances

Donald Norman's account of perceived affordances, which fits in none of the categories presented above, is an attempt to make the concept of affordances applicable for designers. He introduces the notion of perceived affordances to distinguish it from what he calls "real affordances" (Norman 1999). The main difference is that perceived affordances are related to the perceiver alone, in the sense that it is a perceived option for interacting with the environment. Affordances on the other hand "…reflect the possible relationships among actors and objects: they are properties of the world." This rather sketchy account of Norman is not able to spell out what affordances really are in detail, but it might provide an inspiration for understanding how general affordances can become relevant for one perceiver. Interestingly, Norman reintroduces a distinction that ecological psychology sought to overcome: The distinction between the reality in itself and the perceived reality as a mental model. Gibson dismissed both the ideas that there is a genuine, observer independent reality as well as subjective mental representations of that reality, which might deviate completely from 'what there really is'. According to Gibson, ecological psychology and especially the theory of affordances "suggests that the absolute duality of "objective" and "subjective" is false. When we consider the affordances of things, we escape this philosophical dichotomy" (Gibson 1986, 28). This rejection of the traditional philosophical dichotomy is related to Gibson's idea that animals always perceive a meaningful environment in the sense that the meaning of the environment, manifested or given by affordances is already "out there" and nothing to be imposed on a meaningless, objective physical reality. In that respect, Gibson spoke of affordances being objective and subjective at the same time, or neither, rejecting the distinction itself as meaningless. Norman (1999), to adjust the concept of affordances for design purposes, introduces the distinction of a perceived reality and a 'real' reality, namely in his distinction between perceived affordances and real affordances. Unfortunately, as already mentioned above, Norman does not give a substantial account of what the two categories of affordances consist in and what characterizes their difference in detail, but rather adds another problematic description of affordances to the preexisting landscape of accounts and interpretations, with all their flaws and benefits. Furthermore, it is doubtful if ecological psychologists would still consider Norman's notion a proper notion of affordances, as the affordance exists only in the act of perception and thus can best be understood as mentally represented.

3.5.3.2 Experienced Mandates

Siegel (2014) discusses to a special class of perceptual experiences involving the perception of affordances – what she calls 'experienced mandates'. Experienced mandates are "experiences of the environment as compelling you to act in a certain way that is solicited or afforded by the environment" (Siegel, 2014, 2). For Siegel, perceptual experiences have perceptual content, which in turn can only be accounted for in terms of representation. Perceptual experience thus involves representational content (for a detailed discussion, see Siegel 2010).[13] She argues that experienced mandates, and therefore affordances, are represented in perception and are, respectively, part of the represented content of perceptual experiences. However, as experienced mandates "pervade much of our conscious lives, arising both in habitual action and specialized skilled action" (Siegel 2014,

13 Siegel, S. (2010) The Contents of Visual Experience. New York: Oxford University Press.

11), it might be possible that these habitual and skilled actions can be very well described in just dynamic terms without the need for assuming representations at all. This could be the case in experienced mandates that exceed perceptual experiences, i.e., perception of affordances in a situation that become action guiding without having a (conscious) perceptual experience of the afforded properties. There are cases of acting on affordances (e.g., putting a tennis racket back in the bag) that are most often unguided by conscious perceptual experiences but nevertheless purposeful, successful actions. Although cases like these often occur in daily life, this does not entail that there is no correlated perceptual experiences, but only that these experiences remain most often unconscious. In principle, it would still be possible to think about the situation afterwards and recall memorized details of what happened; even these details (i.e., properties) remained unnoticed during acting. This implies that there has been a perceptual state which is, according to Siegel's view, always a contentful, thus representational state. Furthermore, this content will, at least in some experienced mandates, have a rationalizing function: the perceptual content can explain why the subject acted as she did.

Subjects often execute afforded actions instead of merely representing them. Experienced mandates are characterized to have an intrinsic motivating aspect, which should be accounted for by introducing "answerability contents" (Siegel 2014, 21). These are contents which add a motivational aspect to an experience in the sense that the experienced content is in principle also answerable. Answerability in general is given when, e.g., someone hears their name: 'Julia' is generally answerable and is phenomenally different from hearing 'Josie' (cf. Siegel 2014, 6). There is a certain "feeling of answerability" (Siegel 2014, 6) about some experiential content, regardless whether the subject responds to it. The same holds for the contents of experienced mandates: they come with a feeling of answerability, which by itself is determined by various personal, social and moral norms. That a green traffic light solicits street crossing and actually leads to action is due to learned norms. The answerability contents shape the experience in a propositional form and can be expressed by: "It is answered that: X is to-be-phi'd" (Siegel 2014, 24). Siegel argues for experienced mandates as a special class of experienced affordances that are representational and even propositional in nature. An explanation of the afforded behavior always has to be based on the experiential content, which rationalizes the action. To explain the motivation to act on an affordance, propositions that are part of the experiential content can be identified that include answerability contents.

There are some problems in Siegel's account of experienced mandates. First of all, it seems as she presupposes a lot of higher-order cognitive abilities for affordance perception, as all examples and explications of the right kind of contents able to motivate actions have a propositional format. This in turn presupposes the possession and mastery of the concepts the proposition is composed of. Only creatures capable of propositional thought seem able to feel answerability regarding their experiential content. This could be seen as problematic, as it severely limits the applicability of this affordance conception.

Second, Siegel is unclear on the nature of affordances. It seems that her understanding of affordances is that they are perceivable properties represented in experience. This is not Gibson's notion, as he would conceive of perception as direct pickup of information. Furthermore, the role of the perceiver's bodily constitution is not addressed by Siegel, which leaves unexplained why a subject perceives some affordances, but not others. What is missing in Siegel's account is thus a subjective element that explains why some properties that are represented in perception a) represent possible actions and b) play a role for a given subject's actions and goals. Missing these elements, Siegel has to put everything in the actual experiential content: if the content has the properties of answerability and mandates an action, then an affordance is perceptually represented. How the content actually acquires these properties stays unexplained.

Finally, Siegel explains the motivational aspect in represented affordances by a special feel of answerability. Accordingly, experiential content that in addition motivates a subject to act upon that content has to be "answerable". It seems as Sigel is begging the question, by explication motivation to act by "answerability". As I take it, answerability means that in principle, the right kind of content will motivate me to act, whereas content lacking answerability will not motivate me.

It is puzzling why Siegel does not include intentions of the subject in her account, as this would be a less controversial way of explaining why subjects sometimes act upon their experiential content. Seeing the apple in front of her, a subject might reach for it and eat it on one occasion, but not in another, where the subject might not be hungry or suffering from toothache, therefore having no intentions or motivation to eat the apple.

Siegel would have to claim that the apple in the one case is experienced by an experienced mandate with answerability content, whereas in the case where the subject eats the apple, the experienced mandate had answerability content. It is by no means clear that this is the better explanation over an explanation that would consider that perceptual content is by itself neutral, but can play a functional role in different cognitive operations. Perceptual content according to this line of reasoning can give rise to perceptual knowledge, can become part of a judgment or thought, can give rise to belief states or can guide action – by providing the subject with action-relevant information that is complementary to the subject's intentions and motivations.

3.5.3.3 Perception of Q-ability

Nanay's (2011) account of action-oriented perception focuses on the way objects are seen as having action-related properties and argues for a representational account of action-related property perception. The two core claims are: in order to successfully interact with an object one must represent this very object as qualified for the interaction in question, and that the mode of representing an object as qualified for a certain kind of interaction is perceptual. For this purpose, Nanay introduces the notion of Q-ability, which is a relational property implying features of the object and features of the agent – very much in the spirit of Gibsonian affordances. Nanay defends a weaker claim than what he attributes to Gibson, claiming: Gibsonian affordances are best understood in terms of "what we *should* do", whereas Nanay identifies Q-ability "with what we *can* do" with objects (Nanay 2010, 432).[14]

[14] By translating Gibsonian affordances to "what we should do", Nanay seems to overstate the demand character of Gibsonian affordances. Gibson described affordances as invariant and not dependent on the needs of an observer, contrary to Koffka, who defines demand character of objects relative to observer needs (cf. Gibson 1986, 138f).

Q-ability implies that a certain agent perceives features of an object as Q-able, which basically means that an object is Q-able for an agent if the agent *can* Q (with) it. The *can* should not be understood nomologically, but rather in terms of general possibility:

> An object x is Q-able for agent A at time t in circumstances C if and only if there is a sufficiently high number of relatively close possible worlds where A's attempt to Q at t in C succeed. (Nanay 2010, 431)

The main claim concerning Q-ability is that agents represent objects *perceptually* as having the property of being Q-able, such that a tree is represented perceptually as having the property of being climbable and an apple as being edible, always relative to specific agents (squirrels, humans, etc.). Furthermore, Nanay claims that in order to act with respect to an object x, it is a necessary condition to (perceptually) represent object x as Q-able, or in short: "Q-ing x implies representing (not necessarily conscious) x as Q-able" (Nanay 2010, 432). To distinguish actions from mere bodily movements, Nanay defines actions as involving a mental state that precedes the action, that state itself not necessarily being conscious, but definitely representational. The nature of this representation is such that it necessarily involves the goal state of the action:

> […] the general point is that the performance of an action presupposes some kind of representation of the goal this action: of the state of affairs the action aims to bring about. I understand 'goal' to be the *immediate* outcome of the action performed. (Nanay 2010, 434)

Every action such as described presupposes a representation of the desired goal state as a necessary condition. This representation is in most cases non-conscious and can be described in terms of a visual and motoric anticipation of the endpoint of the movements necessary for intentional interaction with an object, such as grasping a cup would involve representing the hand and fingers actually touching or grasping the handle. Apart

---

It is better to interpret Gibson's affordances as offering action instead of demanding for action, as is also implied by the deriving affordances from 'to afford'. The difference Nanay senses is not supported by Gibson's original proposal.

from the goal state, it also involves representing the exact trajectory the hand and arm have to travel to reach the cup, in other words, it involves representing the way the action will be performed (cf. Nanay 2010, 435). From here, Nanay concludes that representation of an action goal (of which the agent is in a position to achieve, to exclude dreaming or fantasizing about impossible actions) necessarily involves representing the object x as Q-able, otherwise the action (with the purpose of achieving a certain goal state) would not even be attempted:

> But I could not represent the way in which I will move my hand there if I did not represent this state of affairs as attainable: if I did not represent the cup as being within my reach: as reachable. In short, performing the action of reaching for the cup implies representing it as reachable. The same argument applies for any other goal-directed action: each time we are Q-ing an object, we must represent it as Q-able. (Nanay 2010, 435)

Having established that objects of goal-directed actions have to be represented as Q-able, Nanay discusses the question whether the representation is perceptual or non-perceptual, i.e., a belief state resulting from prior perceptual states that do not by themselves represent the property of Q-ability but only give rise to the belief state that object x is Q-able.[15] First of all, there is empirical evidence that representing objects in terms of what action they can be used for is at least *one* way of representing objects. A patient with unilateral neglect was better at finding objects that had salient action-related features than finding objects whose primary salient features were visual (i.e., color, shape) and unrelated to possible actions. This finding suggests that visual properties like color or shape are not processed in the same way as the Q-ability properties, as the latter can still be represented if the perception of the former is affected (cf. Nanay 2010, 437). For Nanay, this furthermore suggests that Q-ability is representable in visual perception and not only standardly assumed visual properties like color and shape. The philosophical argument for as to why Q-ability is not represented via a belief-like state but in perception is based

15 Nanay allows for all possible non-perceptual representational states and uses belief state just as one possible example for a non-perceptual state.

on the premise that beliefs imply other beliefs or other mental states in general:

> The non-perceptual account presupposes that we could not have a non-perceptual representation of the object as Q-able unless we had some contextual information or assumption about the object: that it is not going to disappear if I touch it, etc. […] Beliefs are famously sensitive to our other beliefs, but what matters here is something beliefs and any other non-perceptual representations have in common: that they have to be sensitive to contextual information or assumption about the object without which the object would not be represented a Q-able […] We could not have the non-perceptual representation of the object as Q-able unless we had these other mental states. (Nanay 2010, 440)

With this argument, Nanay construes a case where a property of Q-ability is represented against the better knowledge (other contradicting beliefs, assumptions and contextual information) of the agent that the object is actually not Q-able. Unfortunately, the example is not very convincing, as will be shown shortly.

Nanay introduces the case of a person standing behind a Plexiglas wall, knowing that he does so and seeing someone throwing a ball in his direction. He assumes that the subject behind the wall reach out for the ball – in an attempt to catch it, while knowing that this is impossible due to the Plexiglas between them. If the representation of Q-ability were a belief state, this state would conflict with the information that there is a Plexiglas wall and the ball being thus not catchable. This is only a problem if the further premise from the quote above is valid: That non-perceptual representation of Q-ability necessarily involves all sorts of other contextual information without which Q-ability could not be represented. Hence, the information that the ball is not catchable because there is a Plexiglas wall would be necessary for the non-perceptual representation of the ball *as catchable*, which would admittedly be contradicting. From this, Nanay concludes that Q-ability can only be perceptually represented, given that "even if I have all the evidence that the object is not Q-able, I cannot help representing it as Q-able […]" (Nanay 2010, 440). Nanay draws analogy to the Müller-Lyer illusion to make clear how he thinks of the cognitive penetrability of Q-ability perception:

> Perception is famously belief-independent: we cannot help seeing the Müller-Lyer illusion drawing as a picture of two uneven lines even if we know it perfectly well that the two lines are of equal length. Similarly, we cannot help seeing the object as Q-able even if we have all the evidence that it is not Q-able. (Nanay 2010, 440)

There are a couple of problems with this argument. First of all, it is a dubious assertion that the subject in front of the Plexiglas wall will perceive or represent the ball as catchable. Why would he do so, knowing that there is a wall of Plexiglas between them? Maybe the subject cannot help showing some sort of bodily reaction, such as ducking away or jerking a little bit, but it is unclear why the subject should actually attempt to catch the ball in the sense of attempting a proper intentional action. Without providing strong evidence that this should be the case, the example is too weak to support Nanay's strong claim. Second, and more problematic, is the fact that Nanay introduces without further justification that the agent behind the Plexiglas wall cannot help representing the ball as Q-able, which entails the general claim that agents always represent the Q-ability properties objects have. This leads to odd implications about perceiving properties of objects, and the same criticism that applies to Gibson's direct perception of affordances applies here too: In principle, any object has infinitely many affordances or Q-ability properties, but the ones that are actually perceived or represented are those which matter in a given situation, determined, among others, by the subject's intentions and environmental circumstances. It is not clear why the subject in Nanay's example "cannot help" perceiving the ball as catchable, as it would be strange for him to have the intention to catch it, knowing there is a Plexiglas wall in the way, and thus ignore these environmental circumstances. If the subject would actually need to represent the ball as catchable, though knowing that it is not, then the subject would necessarily represent all the other Q-ability properties of the ball too – the ball being throw-able, graspable, bounce-able, roll-able, juggle-able and (infinitely) many more. It becomes even more unlikely when considering more complex affordances: following from Nanay, a subject could never help to see a shoe's laces as tie-able, and the shoes as wear-able, an abacus as doing-the-math-with-able and a horse as ride-able. It is hard to accept why subjects "cannot help" to perceive all these affordances – to some, subjects might just be "blind". It could be that Nanay would be willing to bite the bullet and accept this consequence for the sake of saving his account, but that would certainly not help in rendering the whole account more plausible.

Moreover, it is not even clear what explanatory role Q-ability can play any longer: If the intended actions of a subject should be explained on the basis that the subject perceives or sees Q-ability properties in objects, then perceiving all of the Q-ability properties of an object cannot explain why the subject does Q and not something else - which of the represented Q-ability properties give rise to the action in question? If, on the other hand, only some Q-ability properties are perceived and thus influence further actions, Nanay has to give an explanation why only these and not others do so. And this would most likely involve environmental circumstance and mental states of the subject, such as needs, desires or intentions which would no longer support the claim that the subject in the example cannot help to perceive the catch-ability of the ball although the contextual information and the subject's mental states tell otherwise.

What about the claim then, that Q-ability is perceptually represented and not non-perceptually represented? The whole claim seems to be based on a misunderstanding of the nature of what Nanay calls perceptual representation in contrast to non-perceptual representation. Apparently, Nanay wants to argue for a non-inferential view of Q-ability representation, which distinguishes him from Gibson, whose endeavor was to argue for a non-inferential, non-representational view of affordance perception. It seems that Nanay wants to avoid Gibson's problems with direct perception (see ch. 3.4) and at the same time avoid an "over-intellectualization" of Q-ability representation, such that sophisticated inferential skills, potentially involving conceptual knowledge, is involved in perception of action-possibilities and thus limited to animals of the right cognitive development or developmental stage - which arguably could exclude a lot of animals, primates and human babies from the ability to represent Q-ability. Hence, the real difference Nanay is arguing for is better captured in terms of 'inferential vs. non-inferential' representation of Q-ability. There are two ways to revise Nanay's argument to avoid the consequences just mentioned above, and the solution could well be a combination of the two points: First, accepting the idea that perception is also inferential to some extent would diminish the potential threat of over-intellectualization drastically.[16] Any account of inferential perception that does not rely on beliefs and presupposes conceptual knowledge would do. Second, allowing for representation in terms of possible movements can explain many of the phenomena Nanay mentions without being committed to a non-perceptual account at the same time. For example, Milner and Goodale's (1995, see also ch. 6) two visual systems hypothesis can explain how some environmental features are directly processed in terms of possible movements or actions. In addition, there is evidence that subjects process basic action-related properties even if the they are task-irrelevant, thus being no part of the intentional setting (Ellis & Tucker 2001). These findings have the advantage that they are rather low-level phenomena and hardly involve sophisticated knowledge in terms of beliefs and background information, but rather point to an automatic evaluation of the immediate environment in terms of basic action possibilities.

The difference in this approach to action possibility representation is that not all the Q-ability properties have to be represented and the subject can clearly represent only some in a given situation, while others in another context. If the window is open and a sudden breeze is about to blow my papers away, the water bottle can be represented as a good paper weight, in another situation, the bottle will be a hammer-substitute helping to get a stubborn thumbtack into the wall and in the next context it will be just a container that affords drinking from. All these examples involve representing the bottle as graspable, lift-able and in the thumbtack-case, even as solid. The grasp-ability comes more or less for free, as this seems to be just a basic mode of perceiving our environment in terms of basic actions, whereas lift-ability and solidity are rather likely to involve knowledge in the form of prior experience. I have to know that a glass bottle is solid enough not to break when I attempt to hammer a thumbtack

16 See Hatfield (2002) for an overview of traditional and contemporary views of perception as unconscious inference.

into my wall, and this will by no means be given by looking at the bottle object alone. According to Nanay, the subject would still "see" the bottle as a hammer even if the subject knew the bottle was made of light plastic and only resembling a glass bottle. This is highly implausible, as it is quite obvious that a subject that knows the bottle is not made of glass will not consider the bottle any longer as an option but search for something else. This would be as absurd as assuming that although the tourist in Japan has learned that the wax replicas of the food the restaurants display in their shop windows are actually only a perfect copy of the food and made of inedible wax, the tourist would still represent it as edible. Furthermore, there is other empirical evidence that contradicts the claim that subjects will always perceive Q-ability, even if the circumstances do not allow for Q-ing. Cardellicchio et al. (2013) present evidence that an object's affordances (e.g. of a cup) are only perceived when it is either within reach for the subject or when it is out of reach for the subject but within reaching space for another subject or even an avatar. This suggests that the social situation and other contextual features play a much more important role on affordance perception than indicated by Nanay, who seems to take Q-ability perception as an automatically elicited process as soon as Q-ability is present.

To conclude, there is only a limited range of Q-ability properties that are automatically, probably non-consciously, represented in terms of basic actions, but this is limited to actions based on simple movements, such as reaching and grasping. More abstract actions, such as eating or catching definitely involve knowledge (though not necessarily in terms of conceptual knowledge and beliefs) and thus Nanay's premise that we cannot help but representing Q-ability despite better knowledge is unsustainable. Nanay cannot argue convincingly that Q-ability properties are always represented perceptually – some of them, the rather basic ones in the sense that they are correlated to simple movements, may be represented in Nanay's perceptual way, others clearly involve more contextual information and mental states such as beliefs and even conceptual knowledge. The best way to characterize Q-ability or affordance representations seems to lie in a gradual understanding that allows for an increase in representational and cognitive complexity: simple affordances are correlated with basic movements and can thus be taken to be perceived "automatically", whereas with increasing complexity, more and more cognitive processes are involved in affordance representation so that it hardly makes sense any longer to conceive of it as mere perceptual representation.

3.6 Affordances Represent Possible Actions

From the analyses in the preceding chapters, it can be concluded that Gibson's initial proposal (affordances are objective properties; they are directly picked up) is not viable. Too many problems arise as consequences of Gibson's controversial premises, and although his fellow ecological psychology successors made a lot of effort to save his ideas and avoid these problems, it does not look as if they were successful with their enterprise. The way ecological psychology treats affordances is still problematic and it is hard to see how this concept of affordances can play a major or even central role in any serious psychological science. That said, there is an alternative way of capturing the idea of affordances: Committing to a representational notion of affordances, the troubles that arose from the subjective-objective distinction issues and the direct perception assumption can be avoided – at the cost of having to deal with general problems all representational accounts are facing.

Accordingly, the most viable way to understand affordances is as cognitive representations of action possibilities. Representing action possibilities implies representing features of objects in the subject's environment in terms of possible interactions. Possible interactions in turn are represented as possible movements in terms of sets of potential motor parameters and motor commands. Features of the environment are related to information stored in and retrieved from the body schema. Only features that are commensurate to some extent with information in the body schema are candidates for represented action possibilities at all. This entails that it is highly unlikely that subjects would automatically represent a giant cup as reachable or graspable, and definitely not as something to drink from. All this is true only for simple affordances – more complex affordances have to be understood as representations that are no longer involving specific movements, but represent complex action goals and contextual features which lead to the generation of action. To give an example, a door handle (having the affordance of being graspable and of opening a door) is represented as the exact location of the endpoint of a reaching and grasping movement. This happens almost automatically – whenever a subject perceives a door handle, a possible grasping action is represented, in terms of activating a motor command that would generate the required action. With more complex actions, this direct and automatic connection gets lost, and stored knowledge and acquired skills become increasingly important – to represent the possible affordances of a bicycle, one has to have a lot of previous experience with bikes, such as watching people cycling, or having tried to ride a bike, etc. The connection to specific body parts is no longer given, as riding a bike does not just involve some movements of individual limbs (such as a pointing or reaching movement) but is a complex set of skills and well-adjusted muscular activities, which makes learning how to ride a bike rather difficult. The same holds for the affordance of a lighter to be a bottle opener. Knowledge about levers and the general working of bottle openers has to be (at least in a rudimentary form) available, otherwise it is quite unlikely from the representation of the simple affordances a lighter offers – to be graspable, to fit in one's palms – to conclude that it can be also used for removing a bottle cap. Or consider the affordances of a musical instrument, such as a piano. The expert pianist represents the arrangement of keys no longer as just a series of white and black keys, but as an arrangement of musical scales, chords, tunes and melodies. The novice hardly represents the keys in terms of playing an a-minor chord or playing a c-major scale. Representing higher order affordances of an object requires sophisticated skills that are related to the objects features. The piano only affords sophisticated interaction if the agent has acquired some relevant skills. This implies that only a small set of affordances, the set of simple or basic affordances can be represented automatically and 'directly', and it is only those features of objects that have a clear relation to body parts. Things that are in reaching distance, things that are in accordance with one's grip size, things that correspond to body width or height are among the features that are represented in a simple, body-related way in terms of a 'possible movement format'. Representation of more complex affordances requires a more complex generation of possible movements and actions, which implies retrieval of formerly acquired motor skills, triggered by information processed in the respective context. This process is involved in representing possible affordances of different tools: in order to make sense of a carving tool or wire stripping pliers, one has to learn how to manipulate these tools and what purposes they can be used for. This can happen in a primary or secondary way, either by trial and error, or by observing a skilled user. But without any of these experiences, wire stripping pliers are unlikely to be represented as having the affordance of removing the plastic layer of a cable.

In the remaining chapters, when using the term affordances, I take it to designate action-related properties of objects that are represented by subjects and thus unfolding action opportunities for the subjects. For referring to the ecological properties, I will reserve the term 'Gibsonian affordances'.

In chapter 8, general account of action-related representations will be developed that captures the explanatory value of affordances and other action-related approaches, which will be discussed in the following chapters.

4 Action Representation is Essentially Egocentric

In the previous chapter, I have argued for the existence of very simple action representations, also sometimes referred to simple affordances. These basic action-related representations enable subjects to interact with their environment by representing features of the environment in terms of situated actions. Basic action-related representations are of a very simple structure, so that (flexible) behavior of all kinds of animals can be explained on their basis. In basic action-related representation, features of objects are related to skills of the subject in an implicit way, explaining how subjects with different physical constitutions determine individual action opportunities on the basis of their physical constitution. Being only implicit, the representational content does not require a propositional structure, predicative potential, or possession of concepts.

The central aspect of basic action-related representation is that they are essentially self-related. Representations being able to guide or initiate actions need to involve a reference to the agent to become executable. This idea has been inspired to a great extent by Campbell's (1994) notion of 'causal indexicals', but the idea of implicit self-relation enabling action can be found in various other accounts. The following sections will provide an overview over these accounts before discussing Campbell's idea in more detail. Considering that all the accounts are rather sketchy in nature, I will develop a more substantial account of basic action-related representation maintaining the idea of implicit self-relation but being more explicit about how subjects actually represent action-related features and what their role for action guidance consist in.

4.1 Essential Self-Relation and Action

4.1.1 The Essential Indexical

The starting point for the discussion of implicitly agent-related representation is by introducing Perry's (1979) notion of the 'essential indexical'. Examples for indexical terms are 'I', 'here' or 'now', and their main feature is that indexical terms are context sensitive, i.e., their reference is determined by the respective context of their use or appearance. Perry presents an argument for the claim that some indexicals are essential, in that they cannot be substituted by any other term, e.g. in stating the belief that motivated an action. In an example, a man, John Perry, is shopping in a supermarket and suddenly discovers a trail of sugar on the floor, possibly originating from an open pack of sugar in a customer's shopping cart. He thinks: 'Somebody's making a mess'. Curious to find out who is the perpetrator, he follows the trail to finally discover that the open sugar pack is from his own shopping cart. His belief thus changes from 'somebody's making a mess' to 'I am making a mess', which leads him to rearrange the pack of sugar to stop the sugar bag from spilling (cf. Perry 1979, 3). Crucial to the example is the change in belief that explains the subsequent action. The belief 'somebody is making a mess' was a true belief at the time, as Perry did not think that he was that very somebody who was making a mess. Consequently, it did not lead to rearranging the pack of sugar. This action can only be explained by the belief 'I am making a mess' he came to entertain. The indexical's special role forbids its substitution with another co-referential term, as the action could no longer be explained with the substituted belief, although the truth-value of the belief would remain unaffected. Another term with the same referent would be 'John Perry', and indeed, 'John Perry is making a mess' would still be a true statement. But this statement cannot explain the action, unless a further belief is added, namely 'and I am John Perry', which is why 'John Perry is making a mess' is crucially different from 'I am making a mess' in this scenario. Furthermore, just identifying the belief that explains the action with the sentence 'I am making a mess' is not sufficient, as this sentence can be thought or uttered by any subject. It explains the action of John Perry only in case it is John Perry entertaining the belief. In any other case, if a different subject held this belief, it would be a different belief, a belief that is either not true or cannot explain the action in question or even both. To sum up, there is something special about a certain set of beliefs containing indexicals that lead to actions. The indexical in the beliefs in question is *essential* as only the indexical is able to explain why the subject acted at all – any other description of the belief with an invariable truth value is not able to explain anymore why the subject acted, without presupposing or introducing additional beliefs containing the very same indexical term. There is no non-indexical way to state the belief, so to speak, as we

> [use] sentences with indexicals or relativized propositions to individuate belief states, for the purpose of classifying believers in ways useful for explanation and prediction. (Perry 1979, 18)

Perry concludes that not all belief states can be individuated by propositional content – the propositional content alone would not explain the subsequent action or would be plainly wrong if uttered by a different person, or at a different time or location. There is an essential indexical element in some belief state that is necessary for explaining actions and points to a special relation between the content believed and the state of believe one is in:

> The proposal, then is that there is not an identity, or even an isomorphic correspondences, but only a systematic relationship between the belief states one is in an what one thereby believes. (Perry 1979, 18)

Perry's discussion of the essentiality of indexicals has never intended to be a substantial contribution to philosophy of action or to philosophy of mind. Nevertheless, the notion of essential indexicality can provide useful insights in the nature of action–related representations and could be an addition to the standard belief-desire model of action explanation (cf. Davidson 1967). In this model, actions are understood as the result of a practical reasoning process, having desires and belief states as premises and the planned or executed action as conclusion. If someone prefers mild coffee and believes that milk makes coffee milder, she will, ceteris paribus, add milk to her coffee. According to advocates of the belief-desire model, the action is fully rationalized by the involved beliefs and desires. The propositional content alone is supposed to do the explanatory work.

Perry, however, would argue that at least in specific cases of unique and limited access to belief states, the essential indexical is a core element in the correct description of the agent's belief state and thus completes the explanation (cf. Perry 1979, 19). Even the milk-coffee-case, one could argue, involves an indexical, subjective element that finally leads to action execution and therefore adds up to the explanation: Just having the belief that milk renders a coffee mild and having the general desire of drinking rather mild coffee is not sufficient for pouring milk in one's coffee – it has to be my desire to now have a mild coffee, believing in the feature of milk to make coffee milder, which will lead to pouring milk in my coffee. Accordingly, belief-desire states that motivate actions also need to include a special, irreducible self-relation.

There is a multitude of possible belief states subjects can entertain with infinite many propositional contents. Just in virtue of having these states or being in these states, no actions have to follow from these states and their respective propositional content. Only representational states that include a basic self-relation can have the desired explanatory role in action explanation. According to Vosgerau (2009), every "representation that directly triggers behavior has to refer to the self, the here, and the now, at least implicitly" (Vosgerau 2009, 94). Vosgerau argues that essential indexicals, such as 'I', 'now' and 'here' express exactly this feature of basic action-related representation, corresponding to simple mental representations establishing the essential self-relation (Vosgerau 2009, 94).
Thus, essential indexicality is not merely a feature of linguistic expressions, but an essential feature of action guiding representations. However, not every self-related representation is automatically action-related or being able to guide action. In that sense, the belief that I am 175m tall does normally not lead to any actions.

4.1.2 Self-Relativity Enables Basic-Level Action

In introducing the notion of self-relativity as opposed to genuine self-reference, Smith (1986) develops a similar idea regarding action-related-representation. He claims that self-relativity is something distinct from self-reference but intimately connected, "forming something of a complementary pair" (Smith 1986, 21). Self-relativity consists in the implicit reference to oneself, whereas self-reference is meant to cover explicit reference to one self. Whereas self-relativity enables basic level action for organisms, self-reference in turn enables higher order cognitive operations, such as thought.

Common examples for self-relative expressions are indexical terms. The use of indexical representations is efficient in the sense that one representation with a stable meaning can be used to refer to different objects in different situations. This is efficient because in situations similar to past experiences, irrelevant features can be abstracted from, such as referring to another person as 'you' is sufficient for what many situations demand. Indexical efficiency in this way prevents the subjects from "drowning in details: any facts that are persistent across its experience can be designed out […] and carried by the environment" (Smith 1986, 24). In a similar way, most actions are situated and therefore also context-dependent. The complete meaning of an action cannot be given in terms of the movements involved, but depends, at least in many cases, on the circumstances in which the action occurred. As the same indexical refers to different persons in different circumstances. The same action type can result in different action outcomes relative to the circumstances. Actions are can thus be efficient in analogy to indexicals: for eating different meals, subjects mostly use similar patterns of movements. This kind of efficiency is important, as it would be cognitively exhausting, if individual behavioral patterns had to be developed in every new context. Eating with chopsticks for the first time instead of using the familiar cutlery is such an example.

The circumstantial relativity of action is of interest here because it "requires, among other things, the representation of one's self, because that self is the source of the relativity" (Smith 1986, 26). For a representation to have implications for action at all, the subject of action must be at least an implicit part of the representation. The self-representation needs only to be implicit, which means the representation does not have to contain a part that stands for or refers to the representing subject. An example for such an implicit self-representing could be: 'there's a bear to the right', which implies a subject to which the bear is to the right. The implicit self-representation makes the representation relevant for the agent's life (cf. Smith 1986, 27). Without self-reference, the representation would only represent a detached state of affairs or a property, such as 'Hungry!'. To connect this content to a subject's life and enable action, such as searching for food, self-relativity is necessary and thus a fundamental aspect for all action guiding representations.

4.1.3 Deictic Representations

Agre's (1995) notion of a deictic representation scheme also captures the idea action guiding representations are inherently self-related. His central distinction is between two kinds of intentionality and therefore two kinds of ontologies, deictic and objective, which result in either a deictic representational schema or an objective representational schema. In most representational systems, both kinds of representations can be at work simultaneously, though deictic "intentionality is the predominant form of intentionality in the everyday activities of human beings" (Agre 1997, 243). For Agre, deictic representation should provide an alternative for the explanatorily deficient model-theoretic (computational) accounts of the mind. The deficits become most salient when it comes to explaining agent-world interaction:

> AI research has been based on definite but only partly articulated views about the nature and purpose of representation. Representations in an agent's mind have been understood as models that correspond to the outside world through a systematic mapping. As a result, the meanings of an agent's representations can be determined independently of its current location, attitudes, or goals. Reference has been a marginal concern within this picture, either assimilated to sense or simply posited through the operation of simulated worlds in which symbols automatically connect to their referents. One consequence of this picture is that indexicality has been almost entirely absent from AI research. And the model-theoretic understanding of representational semantics has made it unclear how we might understand the concrete relationships between a representation-owning agent and the environment in which it conducts its activities. (Agre 1997, 241)

In order to explain agent-world interaction, a notion of representation is required that is based on or exhibits some kind of indexicality. Indexicality is an aspect of deictic representation. Deictic representations are part of a deictic ontology, which is "defined only in indexical and functional terms, that is, in relation to an agent's spatial location, social position, or current or typical goals or projects" (Agre 1997, 243). In contrast, an "objective ontology holds that individuals can be defined without reference to any agent's activities or intentional states" (Agre 1997, 243). This distinction is in analogy to the distinction made by egocentric and allocentric frames of reference (c.f. Campbell 1993), in which the position of objects is either defined in relation to a representing subject or defined in terms of the relations the objects have to each other. A map can be understood as an allocentric representation of the position of objects in relation to each other, with no representing subject being involved. The objects of deictic representations are defined as 'entities'[17]: "If an agent has an intentional relationship to an entity then as far as the agent is concerned the latter is defined entirely in terms of the role it plays in the agent's activities" (Agre 1997, 243). Examples for deictic entities are "the-door-I am-opening, the-stop-light-I-am-approaching, the-envelope-I-am-opening" (Agre 1997, 243). These representational or intentional objects are entirely specifiable in indexical and functional terms, as they are specifically related to a subject and have a role in an action of the subject.

Agre claims that deictic representation is more fundamental than objective representation (cf. Agre 1997, 243). Deictic representation plays a major role in everyday interaction with (objects in) the world. Everyday interactions are foremost about opening doors, eating with cutlery, drinking from cups, glasses and bottles, typing on keyboards, etc. What is needed for successful interaction with these objects is functional knowledge in relation to the agent where the function is "indexed" to the agent. Moreover, subjects relate to most objects in everyday life in their generic nature: They treat glasses, stamps and door handles not as individuals, but in their most generic functional being – it is not door handle N°234 I am grasping and pressing down, but *this* door handle that opens

17 Agre uses the term 'individuals' for the objects of objective representation.

this door. What unites most door handles is that they roughly designed the same way and are therefore functionally similar. When an agent adopts an objective intentionality to one of the objects listed as examples, it is most often because problems or some extraordinary circumstances occur. These extraordinary circumstances demand that the agent relates to the object in a more detached, objective way that represents the individual properties of the object that constitute its functional deviance. A door handle which is loose and in danger of breaking apart would require a different treatment than an intact door handle. The almost broken door handle, has to be approached in its individuality. But even in this case, the stored functional knowledge of door handles is at work and enables the agent to successfully interact with the quirky door handle – objective intentionality "is built on top of deictic intentionality as a further complication or refinement" (Agre 1997, 245).

According to Agre, for deictic representation to explain successful interaction of agents and objects, the agent has to represent the functionally significant properties of objects as they play a role for the current interaction context of the agent, instead of representing their general functionality (cf. Agre 1997, 256). Only deictic representations can thus account for spontaneous and dynamic interaction with the environment, as deictic representation involve reference to the subject which in turn determines the detection of context-relevant functional properties:

> The relationships with things that we take up in concrete activity arise equally through our intentions and through our bodily involvement in a physical situation. (Agre 1997, 256)

Although Agre is not specific which kind of perceptual and cognitive abilities are involved in functional property detection, the reference to the subject's body suggests that physical constitution and skills drive functional property selection. This would make sense insofar as not all subjects can act on the same functional properties of objects in the same way, due to physical differences.

4.2 Causal Indexicals – Action-Related Representations Referring to the Self and the World

'Causal indexicals' is a term coined by John Campbell (Campbell 1993, 1994) in the course of developing an account of the prerequisites of self-consciousness. Causal indexicals can be expressed by terms whose reference varies from subject to subject. In analogy to the familiar personal, spatial or temporal indexical terms, such as 'I', 'here', or 'now', they are context sensitive. The difference is that causal indexical terms refer to the 'causal powers' of the subject deploying a causal indexical. Causal powers are determined by abilities and skills of an agent and are related to bodily aspects and learned behavior. Examples for causal indexical terms are: 'this is a weight I can easily lift', 'this is too hot for me to handle', 'this is a gap I can be jump over' or 'this is within reach' (cf. Campbell 1994, 43). Campbell is explicit that causal indexicals are not merely a linguistic phenomenon but are about cognition, thus causal indexical thinking is supposed to be a cognitive mode of representing aspects of one's environment. Indeed, a subject entertaining causal indexical thinking does not need to have all the linguistic concepts involved in expressing causal indexical terms, such as 'I', 'weight' or 'temperature', nor does it need the concepts of a 'self', to think in causal indexical ways: "A creature could use representations of things as within reach or out of reach without having the ability to think using the first person" (Campbell 1994, 44). Causal indexical thinking can therefore figure in behavior explanations of non-linguistic animals such as squirrels, cats, chimpanzees and human infants. It is a non-conceptual or pre-conceptual way of representing interaction possibilities in the world and should apply to every animal capable of flexible behavior.

The most important aspect of causal indexicals is their implications for behavior and actions. As mentioned above, causal indexicals refer to the causal powers of the subject deploying a causal indexical. They are context sensitive and part of the context is determined by the subject. Accordingly, the meaning of causal indexicals changes relative to subjects and their abilities. Hence, what is lift-able for an adult differs from what is lift-able for a much younger child, and the distance of two trees might be perfectly suited to be jumped over for a squirrel, but not for a cat. The same causal indexical can refer to entirely different actions for different subjects. Furthermore, thinking in causal indexical ways has immediate implications for the behavior of subjects – whenever cat entertains the causal indexical 'is too far to jump over', it normally will refrain from jumping.[18]

Causal indexicals yield a cognitive explanation of behavior for a whole variety of animals, notably including non-linguistic animals and pre-linguistic human infants, but also for human adults. Qua being a primitive mode of action-representation, causal indexicals characterize a fundamental representational mechanism regarding simple, everyday interactions. It can best be understood as a primitive subject-world relationship in terms of interaction opportunities that are determined by basic physical features and abilities of the subject. Representing things as within reach is determined by the arm length of a subject and the distance of the object. The causal indexical representation 'within reach' is thus a primitive representation that represents distance of an object in terms of arm length. No concepts of length or distance need to be presupposed for this kind of primitive representation; only a (possible) reaching movement has to be represented towards an object. In representing an object feature in terms of a possible movement, the representation involves an essential self-relation, as the movement in question is the subject's movement, and it is *her reaching* that is an essential part of her representation. At that stage and in these contexts, causal indexicals can always only be representations of the subject's own possible movements in relation to objects.

As mentioned above, Campbell's notion, together with the other self-relativity accounts discussed earlier, is rather sketchy. To develop these ideas further, some aspects have to be clarified and elaborated. This discussion will result in a more substantial notion I want to call 'basic action-related representation', to avoid confusion with any of the other accounts discussed in this book.

18 For simplification, it is assumed that normal circumstances hold, without threats such as being hunted by predators, or without perceptual disturbances, etc.

First, basic action related representation represents features of the environment in terms of possible movements. The format of the representation is best characterized as 'movement'.

Second, the representation of both objects and the subject is implicit, allowing the structure of the representation to be as simple as possible.

Third, basic action-related representations crucially involve information of the subject's body, which is provided by integrating information stored and processed in the body schema.

4.2.1 Representations in a Movement Format

Causal indexicals are supposed to represent features of the environment in a simple, self-related way, so that no demanding cognitive resources have to be presupposed. Such a simple mode of representation could consist in a movement format, which describes a format that is comprises elements such as motor parameters, motor patterns and motor commands. To elaborate this idea, I will present accounts that consider action representation an important aspect in representing environmental features, such as spatial and object representation.

To start with, let's consider an everyday example to illustrate what could be meant by movement format: Most often, people enter their numerical cash card PIN when withdrawing money on an ATM's keyboard of a certain layout. Assumed that subjects are normally exposed to the same keyboard layout, after a few times of entering the PIN, it becomes an almost automatic process and most subjects do no longer have to consciously remember the PIN and then press the keys, but rather "let their fingers do the work". A movement pattern is stored in addition to the numbers. Over time, it is even likely that subjects take longer merely mentally recalling the numbers than by simply typing the PIN. Everybody who has been confronted with a unfamiliar keyboard layout (e.g., ATM keyboards in Japan do not have a block layout like the European one's, but feature a horizontal row of numbers) knows about the initial irritation when trying to enter the PIN – it can take a while "translating" the stored number representation into the new format required. Cases like this suggest that the process of storing and retrieving of information happens not only in a symbolic or imagistic way, such as picturing the numbers as printed on a letter from the bank, but crucially involves the stored pattern of finger movement. Having a movement pattern "at hand" is a quick, simple and efficient way of retrieving this kind of information that works for many situations.

This example can be interpreted as case of storing information in a movement format. The information becomes associated with a movement pattern that can be further specified in terms of motor parameters and motor commands, and can even become the standard mode of representing the information. There is empirical evidence which suggests "an important role of body actions in arithmetic processing" (Tschentscher et al. 2012, 3140). Studies on the influence of forced gestures and finger movements in arithmetical tasks showed that children using gestures and finger movements during arithmetic task had a better problem solving performance than children in the control group who were not allowed to produce any gestures, furthermore it had an effect on acquiring new theoretical mathematical knowledge later on (c.f. Tschentscher et al. 2012; for an overview to embodied numerical cognition, see Fischer 2012). Although there is no direct evidence that information like a four digit PIN is actually stored in a motoric format, this would explain why subjects sometimes have problems retrieving a combination of numbers in the abstract realm when a typing opportunity is not given. A similar example can be construed by considering expert musician, such as piano players or violinists, who understand sheet music in terms of the finger patterns they would use for playing. Of course, expert musician can "read" music and most likely will have an auditory representation of the score's content. However, they will also directly translate musical notes into movements, using stored motor patterns acquired over time and strongly associated with the sound produced and perceived.

4.2.1.1 The desert ant's odometer

A specific case of a representation in a movement format is the way desert ants represent the distance to their nest (cf. Vosgerau 2009, Wang and Spelke 2002, Wittlinger et al. 2006). Desert ants typically have an unsystematic foraging behavior, meaning that they show random search behavior until they find food, which causes them to return to the nest on a direct path. How do the ants know what the shortest distance to their nest is?

They must have some sort of navigational device or mechanism that tells them exactly where they are once they found food, and in which direction and how far they have to walk to carry the food back to the nest. Landmark navigation cannot explain this behavior, as there are hardly any landmarks in desert areas for desert ants to exploit for navigational purposes. In addition, ants that are on their way back to the nest and are moved to another location will continue their way parallel to their original path. They would have been disturbed by the new environmental layout if landmark navigation would be the underlying mechanism (cf. Wang and Spelke 2002, 376). How do desert ants represent the location of their nest? The direction in which they have to walk is explained in terms of a constant calculation of the angle of the sun to the ant and the nest (cf. Gallistel 1993; Wang & Spelke 2002). Thus, the representation of the ant's current location is based on a process involving dynamically recalculating the ant's position. This mechanism works similar to how modern navigation devices function, instead of GPS-satellites, the ant uses the sun as celestial point of reference. The distance the ant has to walk is represented by a number of steps the ant has to walk. Wittlinger et al. (2006) found that ants have an inbuilt mechanism they call the 'ant odometer', which has the function of counting steps. Their experiments involved two groups of ants, whose legs had either been shortened, or elongated with tiny stilts. The result was that the ants with shortened legs stopped their homing behavior before they reached the nest, whereas ants with elongated legs walked past the nest. These findings show that the ants walk a precise number of steps, which (under normal circumstances) would represent the shortest path to the nest. To accomplish this, the ants must dynamically calculate the amount of steps needed to return home to the nest, relative to their current location. The example of the ant odometer is evidence for the existence of a movement format or representation. The ants represent distances in terms of movements they have to execute. Distance thus means a certain number of steps for the ant.

The ant's representation is causally indexical. A feature of the environment (the location of the nest) is represented in terms of possible movements. Neither the location of the nest nor the ant agent are explicitly represented, at least in the ant's case this is more than unlikely: nobody would like to ascribe a predicational capacity to the ant, neither the possession of a self-concept. Nevertheless, the ant's representation is implicitly representing the location of the nest, as its only purpose is enabling the ant to return home. As the number of steps representing the location is the steps the ant entertaining the representation has to walk, the representation is also implicitly self-related: the ant represents its own movements. In that sense, the behavior of a squirrel successfully jumping from tree to tree can be explained by recurring to causally indexical representations that represent distances of trees as either jumpable (within reach for this very squirrel) or not jumpable (i.e., out of reach).

4.2.1.2 Egocentric space is action space

Evans introduced in 'Varieties of Reference' (1982) – and even more prominently in 'Molyneux' Problem' (1985) – the idea that spatial representational content can be accounted for in terms of possible behavior. Thus, egocentric spatial representations get their significance by behavioral possibilities applied to the represented space. An 'up' or 'down' representation gets its significance for a subject by being related to possible behavioral options, and not by reference to body parts: 'down' cannot derive its meaning form 'where the feet are', because it would lose its meaning when the subject finds herself upside down. Egocentric spatial representations thus derive their meaning from the behavioral space of the subject:

> We envisage specification like this: he hears the sound up, or down, to the right or to the left, in front or behind, or over there. It is clear that these terms are egocentric terms; they involve the specification of the position of the sound in relation to the observer's own body. But these egocentric terms derive their meaning from their (complicated) connections with the actions of the subject. (Evans 1985, 384)

Spatial representations are expressed by a "vocabulary, whose terms derive their meaning from being linked with bodily action" (Evans 1985, 385). Although the content of the spatial representations is expressed by a vocabulary linked to the actions of the subject, it can never be reduced to specific types of behavior. 'To the left' thus can relate to all sorts of possible actions in the behavioral space, such as grasping, reaching, jumping etc.; the meanings are not derived in the sense that types of representations can be (intimately, reducibly) linked to types of behavior. Evans' idea of accounting for spatial representational content in terms of the subject's action space is inspired by Poincaré's (1958) notion of representative space:

> To localize an object simply means to represent to oneself the movements that would be necessary to reach it. It is not a question representing the movements themselves in space, but solely of representing to oneself the muscular sensations which accompany these movements and which do not presuppose the existence of space. (Poincaré 1958, 47)

Poincaré claims that spatial representation can be given solely in terms of muscular movements, while representing these very movements does not presuppose the existence of space – i.e., the existence of spatial concepts. Evans and Poincaré share the idea that action is central to the representation of egocentric or behavioral space. Assuming that representation of space is fundamental for representing one's environment, action plays a fundamental role in the way subjects perceptually represent their surroundings, which is in terms of possible movements. Egocentric space is thus a product of combining sensory input information with possible actions of the subject:

> Auditory input, or rather the complex property of auditory input which codes the direction of the sound, acquires a spatial *content* for an organism by being linked with behavioral output. (Evans 1985, 385)

Of special importance is Evans' insistence that processing perceptual stimuli is not the result of any kind of inference, but rather that the information is immediately available in a format that allows for acting upon the information:

> We do not hear a sound as coming from a certain direction and then have to *think* or *calculate* which way to turn our heads to look for the source of the sound etc. If this were so, then it should be possible for two people to hear the sound as coming from the same place

> ('having the same position in the auditory field') and yet be disposed to do quite different things in reacting to the sound. Since this does not appear to make sense, we must say that having the perceptual information at least partly consists in being disposed to do various things [...]. (Evans 1985, 383)

In this respect, spatial perception is closely linked with possible behavior, at least in this basic sense. But does this entail that the connection of spatial representation and behavioral output is a foundational one, or could one argue that this is just learned behavior that is triggered on certain occasions? It makes sense to assume that the representation of egocentric space is the most basic form of spatial representation from both a phylogenetic and ontogenetic developmental point of view. Granting this, all forms of detached, allocentric and abstract spatial representations are likely to be grounded in these primitive forms of spatial behavior. Starting from perceiving one's own behavioral space of reachable objects, at later stages of development, one is able to accounts for reachable objects for other people by an extending one's frame of reference from purely egocentric towards a more objective stance. Finally, a general representation of distance and relations of objects in the world can be developed on these grounds, without reference to one's own possible actions. Evans notion of spatial representational content in terms of action space explains how subjects develop spatial representations at first place and allows for further development of more detached spatial representations.

4.2.1.3 The theory of common event coding

The theory of event coding (TEC) (Hommel et al 2001; Hommel 2004; Hommel 2009) also provides as sense for the idea that causal indexicals are representations in a movement, action related format. TEC claims that perceived features of the environment are encoded in the same format as action plans. At the core of TEC is the rejection of the traditional and common "assumption that perceiving a stimulus object and planning a voluntary action are distinct processes operating on completely different codes" (Hommel et al. 2001, 860). Instead, their central claim is that perceiving and action planning is functionally the same, as they both represent external events (Hommel et al. 2001, 860). This is based on the view that perception is both involving and enabling actions, as well as goal-directed action presupposes and generates perceptual input. Therefore, it seems justified to assume that perception and action share a common representational format, rather than assuming complex translation processes between these two domains. Moreover, TEC is based on findings that action is represented quite similarly to visual object representation, which suggests that "principles underlying the organization of perception and action related information should be comparable" (Hommel et al. 2001, 861).

Furthermore, TEC makes a distinction between distal and proximal information. Proximal information is given in terms of sensory and motoric systems, whereas distal information is given in terms of feature codes in the common-coding system. Both representations underlying action planning and object perception share a cognitive code, as they are about distal events. This means that what is represented by the common code is the object of the perception or the action plan regarding that object – both already on a higher level of abstraction. The proximal features are still represented in a domain specific code. Thus, what is represented when, e.g., a cup in front of a subject, is the action of grasping it, and not the proximal motoric activation or muscular movements. But although this is not part of the common code, the proximal information is still related to the commonly coded representation (cf. Hommel et al. 2001, 862 Figure 1). Though being about representing perceptual objects in a code that is at the same time an action code, it seems that TEC does not allow for the most basic causal indexicals. TEC entails the representation of objects and events in an abstracted common code, instead in terms of simple movements, as causal indexical are described. Thus, with TEC, an adequate explanation of the behavior of animals that are cognitively less sophisticated might not be possible, and would therefore diminish the explanatory power of causal indexicals. However, TEC seems like a good model for explaining the next steps in cognitive development, as soon as the first more abstracted representations emerge from the basic level ones. The common code is still strongly associated with perceptual and motoric functions and processes, thus applying to causal indexicals that are about to become more generalized representation. In chapter 9 on abstraction, this point will be further elaborated.

4.2.1.4 B-formats and B-contents

Goldman and de Vignemont (2009) introduce the idea of bodily representations that are primarily individuated by their format and not by their content. They distinguish between bodily *content* and *bodily* format. A mental representation can be about the body, thus having bodily content, but the representational format could yet be amodal or propositional, as in 'that my legs are crossed'. Regardless of content, mental representations can also have different formats, such as visual, auditory, conceptual, or a bodily format:

> A motoric format is used in giving action instructions to one's hands, feet, mouth and other effectors. A somatosensory format represents events occurring at the body's surface. Affective and interoceptive representations plausibly have distinctive B-formats, associated with the physiological conditions of the body, such as pain, temperature, itch, muscular and visceral sensations, vasomotor activity, hunger and thirst. (Goldman and de Vignemont 2009, 3)

The different formats can still have overlapping contents, as a given content can be multi-modally represented, so the content is not what defines bodily format. Gallese and Sinigaglia (2011) elaborate further on this notion and state that representational format constrains what a representation can represent. (cf. Gallese & Sinigaglia 2011, 513f). Thus, a representation is of a bodily format, when the constraining factors of what can be represented are bodily factors. Bodily factors in turn are facts of the body, such as arm-length, strength, hand span, size, posture, etc. A goal representation is of a bodily format, when bodily factors determine the possibility that the goal state obtains. For instance, grasping a cup is represented in a bodily format because the represented goal corresponds to bodily factors. In contrast, an oversized cup of 2 meters height cannot be represented as being graspable in a bodily format, but only in a non-bodily format (e.g., propositional), because the relevant motor program is not available, as there can be no motor program for a reaching and grasping movement that is far beyond a subjects body scale. Whenever a representation represents something in terms of bodily aspects of the representing subject, it is of a bodily format. A cup that is represented in a bodily format is represented in terms of a reaching movement. Thus, only the aspects of the cup relevant for the reaching action will be represented, such as the object size in relation to hand size, object's distance in relation to posture and reaching length, etc. The same cup, in a propositional format will be represented in terms of conceptual constraints: that it is made of ceramics, for drinking tea, is red etc. None of these represented aspects bear (an immediate) relation to bodily factors, thus cannot be constraint by them.

The bodily format of a representation can account for the action-relation in causal indexicals. 'Within reach' and 'is jumpable' are to examples for representational content is entirely constraint by bodily factors. Such a bodily format can also be the movement format described above. The ant's representation of the location of the nest is constrained by the execution of a motor pattern that allows only for specification in limited dimensions. The representation of the ant will always be of the form 'n-steps after turning by α', and thus be a possible movement representing an environmental feature.

4.2.2 Implicit Representation of the Agent and Environmental Features

Causal indexicals are defined to be cognitively primitive, implying that their structure is a simple as possible. For providing an adequate explanation of the behavior of a wide range of animals, causal indexicals need to be able to represent some aspects of the environment while being self-related in a minimal sense. Furthermore, abilities such as conceptual thinking or property predication would limit the applicability severely and lead to an 'over-intellectualization' of the behavior explanation, which should be overcome by referring to basic action-related representations such as causal indexicals. Accordingly, the only way of representing features and self-relation that satisfies these constraints is implicit representation.

The aspect of implicit self-relation is provided by format of causal indexicals. By being representations in a movement format, which represents features in terms of possible movements, it is implied that the movements are the subject's movements. 'Within reach' implies that something is within reach for *me,* if something is 'too hot to handle', it is a representation that can have only immediate implications for the representing subject's actions. In addition, no concepts are needed for these kinds of implicit self-relations. A subject approaching a hot pot and sensing the heat will immediately withdraw the reaching hand. Obviously, the subject was representing the pot as 'too hot' for herself, otherwise the behavior would make no sense. However, the causal indexical 'too hot' is not ascribed by the subject to herself, but already given in being the bearer of phenomenal experience. The self is always already implied by being the subject that detects features. In that sense, it does not make a difference if a non-linguistic toddler or an adult represents something as too hot, as both will need no explicit 'self' ascription for representing the feature of something being 'too hot' for themselves.

The other implicitly represented aspect concerns the features of the environment. To be implicitly represented in this case means that no property is attributed to an object. Thus, a causal indexical of 'within reach' is understood to ascribe no property of 'being within reach' to an object, such as a bottle on the desk. Rather, the causal indexical expresses the presence of an action possibility towards an object that is solely specified by the action possibility. Causal indexical representation thus means that, in a given situation, an object is referred to according to the action possibilities it allows for. Again, no conceptual knowledge is necessary to refer to an object as 'within reach' in terms of a possible grasping movement. The object can be reached for regardless if it is represented as falling under a category. What is crucial is the mapping of possible action onto objects in the behavioral space. These simple action involved in causal indexicals do also not need to be conceptualized to be executable. Reaching, grasping, jumping and pointing are all actions that are intimately connected to the bodily constitution and comprise the set of basic movements from which many complex actions are composed. Reaching as opposed to knitting is learned at the very early stage of development, whereas knitting is an action that demands sophisticated motor control and a rather abstract goal representation, therefore being a conceptualized action that involves other actions. To represent a set of knitting needles as 'is a thing I can knit a scarf with' is involving a huge net of concepts and skills and is no causal indexical, but an explicit and conceptual action-representation.

4.2.3 The Role of the Body Schema

A prominent notion in accounting for the type of body-related information that is enabling and guiding behavior is the concept of a body schema. The body schema crucially provides relevant information for causal indexical representations. The concept of schemata for the control of movement, stored in the sensorimotor cortex has already been mentioned by 19th century neurologists (for an overview, see Gallagher, 2009). The notion of a body schema has been developed further by Head (1920) and Merleau-Ponty (1945/2012; see ch. 2) and others throughout the 20^{th} century, while nowadays most prominently playing a major role in the work of Gallagher (1995; 2005), who proposes a conceptual analysis yielding a systematic distinction between the often confused notion of body image and body schema. Gallagher, approaching embodiment from a phenomenological perspective, claims that higher-level cognitive phenomena, such as phenomenal consciousness and intentionality is grounded in operation of the body image and the body schema, the latter enabling posture maintenance and movement.

According to Gallagher (1995),

> a body schema involves an extraintentional operation carried out prior to or outside of intentional awareness. Although it has an effect on conscious experience, it may be best to characterize it, as Head did, as a subconscious system, produced by various neurological processes, that play an active role in monitoring and governing posture and movement. […] Even in intentional bodily motion, certain postural adjustments of the body that serve to maintain balance are not under conscious control. Various muscle groups make automatic schematic adjustments that I remain unaware of […]. The body schema […] functions in a holistic way. A slight change in posture, for ex-ample, involves a global adjustment across a large number of muscle systems. (Gallagher 1995, 228f)

Gallagher mentions two important distinctive features of the body schema: It is a subconscious, neurological system relevant for guiding posture and movements and holistic in that the information stored and processed on the body schema is subject to global change and therefore involved in all possible movements, as well as all possible movements affect and make use of the body schema. The body schema allows for successful and accurate movements even in cases where the whole awareness id focus on something completely different – such as having a deep conversation while wandering through a complex building, opening and closing doors, using steps and the like. The body schema enables this by means of "operating in a tacitly lived (nonobjective) space, automatically [taking] measure of its environment" (Gallagher 1995, 230). A subject's performance would be worse in general if the subject had to rely on conscious cognitive processing of body parts and posture for action guidance and movement control. The subject would have to think about individual movements, such as the steps involved in walking across the room to switch the light on. The body schema enables this to happen automatically, without having to consciously attend to the specific movements involved in everyday actions.

The ability to execute movements unattended therefore needs to be based on subconscious processes which make use of a constantly updated, dynamic representation of limb position and general posture, as well as general and implicit information about the body, such as size, width, arm and leg length (cf. Longo & Haggard 2010). The information is implicit in the sense that does not represent the objective body size of 1.75m, but rather a relative height relating the subject's body to the environment. Accordingly, the posture and location of the limb at any given time are not represented absolutely, but relative to other body parts and objects that are involved in motor goals.

In the discussion of the deafferentiated patient IW, it becomes clear that one primary source for the body schema is proprioceptive information (from kinetic, muscular, articular, and cutaneous sources), although other sources of information exist (Gallagher & Cole 1995). With IW losing his entire proprioception and tactile feedback, he also lost complete control over movement and posture and had to learn to control his movements from scratch. IW is able to walk and write and even drive a car again, but needs to consciously and visually attend to and guide his movements – when IW is in a dark room and can no longer see the position of his limbs, he is out of control as he completely lacks proprioceptive information about his posture. In contrast to the body schema, Gallagher describes the body image as being a complex set of intentional states, whose intentional object is the body. The intentional states involve perceptions, mental representations, beliefs, etc. (Gallagher 1995**)**. For controlling his movements, IW thus uses the body image substituting the missing information form the body schema. This suggests that one of the major informational sources for the body schema is proprioception, as the body schema enables movement control and posture, which both broke down as the proprioceptive information got lost. IW movements never got automatic again, which also implies that the body image is a less permanent source for movement control, but will guide movement while visual attention is given. The distinction that the body image equals conscious, reflective intentionality, whereas the body schema consists of unconscious, subpersonal processes is not definite. Gallagher does allow for certain, limited interactions of the two systems. Thus, as demonstrated by IW's case, the body image can take over some of the functions of the body schema, although it will always be limited and never function in the same way.

For the present discussion about the specification of body-related information for causal indexical representation, the concept of the body schema and its implications for movement control seems to be the adequate structural entity. Besides kinetic and muscular proprioception, visual proprioception is an important source for the body schema (Gallagher and Cole 1995). Another important aspect of the body schema involves motor habits in the sense of learned or innate movement patterns. Examples that involve motor habits are: swimming, walking, writing, swallowing, etc., where some skill acquisition, such as learning to swim or to write involves a higher degree of conscious attention thus information from the body image. Once these skills are acquired, they become (almost) automatic movement patterns that are executed without consciously attending to. In all these cases, proprioception nevertheless plays an important, maybe even constitutional role for successfully executing the stored movement patterns, as constantly updated information about the position of the body parts is required. In that sense, motor habits are intimately connected to proprioceptive information, in both acquiring and executing the very movement patters.

In addition to Gallagher's (1995) body schema, a 'body model' has been proposed by Longo et al. (2010) to account for the representation of body metric or body size: the body model. They claim that "locating body parts in space requires a combination of afferent information and stored representations of the body" (Longo et al. 2010, 12), thereby arguing against Gallagher that only proprioceptive feedback is not yet enabling actions. The main argument is that no "afferent information provides such information about body size" (Longo et al. 2010, 12). They conclude that an innate body representation exists, the body model, which is supposed to interact with the postural model (i.e., body schema), and thereby provides the desired information of both limb posture and size. Important to mention is that the body model is supposed to be innate and lacks a genuine input channel.

This lack of input channel, however, is deemed highly problematic by Cardinali (2011), who agrees with Longo et al. that the posture information has to be integrated with size information, but refrains from positing a new kind of body representation. The body model should have explanatory value. However, conceiving of the body model as innate renders it implausible - if information about one's body is innate, how would it account for change? Especially implausible is the conception of a body model that does not receive sensory input - without sensory input it could not be expected to provide adequate metric information about the body, which undergoes rapid and constant change (cf. Cardinali 2011, 56ff.). Cardinali comes to conclude that

> we should have a representation that is innate (while body size changes can be influenced by many external environmental factors), "unfed" (that makes difficult to understand how it can be updated), dramatically distorted (that make difficult to understand how we can perform accurately any of our daily motor actions) and unable to follow normal changes in size like, the growth of our own body. It is, indeed, quite difficult to agree on the need of a BM [body model; T.S.]. (Cardinali 2011, 59)

Her alternative suggestion is extending the notion of the body schema, to represent size and dynamically update changes, e.g., in cases of tool use where tools extend the reaching distance in healthy subjects.

In Cardinali et al. (2009), the main hypothesis is that if tool use changes the subject's actions after a training phase with the tool in question, the impact on performance can only be explained by changes in the body schema, which is directly subserving action. The tools used in the experiment extended the subject's reaching distance, which was the only modulating factor the tool provided. They tested healthy subjects if there is any measurable change in performing grasping and pointing movements before and after a training phase of 15 minutes with a 40 cm grabber tool, with which the subjects had to grasp an object. Movement time and acceleration peaks was measured before and after the training phase and found latencies in post-tool-use gasping and pointing movements as well as a decrease in acceleration peaks. The results support the claim that tool use induces a morphological change in the body schema (cf. Cardinali et al. 2009, 479). This suggests an interpretation that the body schema was adapted to the altered limb size, and is therefore able to represent body size in general. The altered body schema represented a different arm length, which in turns changed the whole movement succession (measured with kinematic analysis). The grasping movement of the hand was not affected, which leads to the interpretation that only the arm size is represented differently after tool use. In a control experiment, it could be ruled out that the spatial representation of the object position alone was subject to change and not the body schema, by measuring accuracy in pointing to stimulated points on the arm with the untrained hand before and after tool use. The differences found supported again the interpretation that the arm size representation changed in the body schema (cf. Cardinali et al. 2009; Cardinali et al. 2012).

These findings are important for two reasons: First, they demonstrate the plasticity and dynamicity of the body schema, which is not only able to update posture, but also size variation. Second, this supports the interpretation that in the tool use paradigm, the change of action space representation is subject to a change in the body schema and not vice versa:

there is a primacy of the body schema that determines spatial representation:

> Space representation depends on action possibilities: the representation of what is near (or far) is built on the fact that I can (or cannot) act in that particular region of space. This action supremacy is of great importance: if we push the reasoning further we can readily realize that the decision about the possibility of acting on a particular region of space depends in turn on the knowledge about the size of the body that, in an action context, might be provided by the BS [body schema; T.S.]. (Cardinali 2011, 67).

Accordingly, the body schema determines the possible space of interaction, which is also known peripersonal space, the space surrounding a subject in which it can immediately act. This is further empirical support of the claim that causal indexicals represent (features of) objects in terms of possible movements, by exploiting information stored in the body schema. A causal indexical, such as 'within reach', thus represents the location of an object in terms of the distance one has to reach out, in order to grasp this object.

Body-related information, such as size information or other basic properties, such as one's weight or relative strength are all aspects of the body-relation of causal indexicals that (visual) proprioception can account for. A subject of a given weight will receive different feedback from the different surfaces, textures and substances one eventually encounters in life, and associate this with different actions and states, e.g. when walking, sitting or lying down. A sense of one's own body height can evolve, e.g., on the basis of relation from eye height to invariant structures in the environment and the sensorimotor contingencies involved (cf. Proffitt & Linkenauger 2013). At the same time, parts of the body are always part of the visual field, and thus visual information that is mainly processed unconsciously provides information about relation to other objects and the bodily dimensions. The body-related information should be understood to be of rather simple structure, primarily with the function of enabling basic movements and interactions and not for propositional thinking about one's body.

To sum up, the notion of the body schema as defined by Gallagher (1995) and others provides an understanding of how the body-related information necessary for movement and action is represented. It is a subconscious, subpersonal, almost autonomous and automatic informational organization, which is enables posture and movement and contains other basic aspects of the body. Its main source is proprioceptive information that dynamically updates the information in the body schema, regarding posture, body dimensions, movement patterns and action skills. Moreover, the body schema is also able to integrate and process information of tools and other external objects attached to the body, therefore changing and adapting the action relevant body dimensions. The last point is of some importance, as it explains a common phenomenon: If subjects routinely interact with specific objects, they will be integrated in the body schema consists. That change of the body schema allows for different or new actions and movements, which explains one of the foundations of skill acquisition. For example, musicians and craftsmen, quite often claim that their tools or instrument literally feel like a part of the body, which means that the instrument literally has become a part of the *body schema.*

4.2.4 Developmental Aspects of Causal Indexicals and the Body Schema

If the body schema is the locus of body related information, enabling movement, which in turn is the representational prerequisite of causal indexical representation, the body schema has to come into existence prior or simultaneously to the ability of causal indexical representation. In addition, as causal indexicals are supposed to be developmental simple, the body schema also has to exist from the very early stages of development. There is some dispute whether the body schema is innate or at least established prenatally via early proprioceptive information in the mother's womb, or if it is developed postnatal not before the third to sixth month. The latter view is credited to Merleau-Ponty (1945/2012; see also Gallagher & Meltzoff 1996), whereas more recent findings suggest that the body schema is best conceived of being innate, as, e.g. newborns already show imitation behavior that can only be explained with subpersonal processes enabling motor control in accordance to perceptual stimuli (Gallagher & Melzoff 1996, 212). What is of importance for a definition of causal indexicals is that both parties (innate vs acquired) agree that the body schema exists from very early on. Either the body schema is truly innate, or it "functions as if it were an 'innate complex' [...] that is, as strongly and pervasively as if it were innate, but, as an acquired habit with a developmental history, it is not innate" (Gallagher & Melzoff 1996, 213). From this, it is safe to conclude the body schema exists from very early on and can therefore be associated with perceptual input, enabling causal indexical representation.

4.3 Summary

Causal indexicals are representations of features of the environment in terms of possible actions or movements. Causal indexicals essentially self-related, by being about movements of the subjects, thus do not explicitly represent neither the subject nor the environmental feature or object referred to. The meaning of a causal indexical is determined by the individual subject's 'causal powers', consisting of bodily aspects, skills, acquired movement patterns, etc. Furthermore, causal indexicals are basic representations of primitive structure, being non-conceptual and thus do not presuppose sophisticated cognitive abilities, such as concept possession. The representational format, being of a simple structure, consists in a movement format. Possible movement representations, involving motor plans, motor patterns and motor commands encode environmental features and enable the subject to immediately act upon the detection of the very features. Thus, casual indexical have direct implications for action and are able to explain the action of a broad spectrum of animals. Different accounts exist that can be used to get a better understanding what representing in a movement format means, showing on different levels of complexity the direct involvement of movement representation in cognitive operations. Examples such as of the desert ants' navigation skills in terms of step counting provide strong evidence for a movement format and its function. The idea of bodily formats focuses on the bodily features that determine possible representational contents. Bodily formats are only one way of representing a given content, thus allowing for co-existence of different representational formats. This entails that the bodily format is more fundamental than, e.g., a propositional way of representing a given content. Finally, the notion of 'body schema' can be used to account for the informational resource that determines the range of possible movements in causal indexical representations.

From a developmental perspective, causal indexical representations are one of the most basic cognitive representations, as this way cognitively interpreting one's environment in terms of possible movements can be attributed to all animals that show flexible behavior. Even in the ant's case, similar, however limited, representational abilities can be described on the basis of the ant's behavior. Causal indexicals, as analyzed in this chapter, play a foundational role for explaining animal-environment interactions and, in addition, for the development of more complex representations in ontogenetic development. Animals that are disposed to undergo cognitive development, will built upon the basic self-relation that is crucial to causal indexicals to develop more detached, abstract representations. A detailed discussion of the possible abstraction mechanisms on the basis of casual indexicals and other action-related representations will follow in chapter 9. In the following chapters, I will use the term 'basic action-related representation' when I want to refer to the elementary representations such as causal indexicals.

5 Action-Guiding Representations

In this chapter, I will discuss accounts of action-related representation that share the central claim that mental representation is best understood regarding its function for the cognitive system. The primary function of representation is to enable and guide action, all other functions representations can have in a cognitive system are derived from the original function. These accounts are interesting for the purpose of finding a general account of action-related representation, as the action-guidance accounts approach the topic from an evolutionary perspective. The main premise is thus that representational systems were advantageous from an evolutionary stance because they allowed for more flexible behavior. Once behavior is representation driven, the mechanism that initiates behavior becomes decoupled from direct stimulus detection. In non-representational systems, a given stimulus will normally cause a determined behavioral response, whereas in representational systems, different stimuli can cause the relevant behavior and different behavioral responses to on stimulus type are possible. This expands the behavioral flexibility significantly, as, e.g., in changing environmental circumstances, the mechanism responsible for stimulus processing and triggering behavior can adapt to new stimuli. The dimension of functional adaption and evolutionary selection advantage will become an important aspect in the general account of action-related representation. Representing one's environment in terms of different action possibilities is able to give a cognitively adequate explanation of the behavior of many species. Moreover, it can explain how a mechanism that allows for flexible behavior in relation to possible actions, determined by the physical constitution of an animal, is of evolutionary advantage. Furthermore, if representational systems have evolved on the basis of action-guidance, the development of other, more abstract (mental) actions can also be accounted for on the same basis of action-related, action-guiding representations.

5.1 Pushmi-Pullyu Representations

Millikan's (1995) account of pushmi-pullyu representations (PPRs) has been developed with the background of her general account of an evolution based 'biosemantics' (cf. Millikan 1994). The theory of biosemantics holds that the content of representations is best addressed in terms of consumer-system based proper functions that have evolved naturally over the course of time. Thus, mental representations have a meaning for a consumer system in term of a naturally evolved proper function. To take one of Millikan's (1994) own examples, the splashing of a beaver with his tail produces a representation because other beavers, the consumers, interpret the tail splashing as danger and act accordingly. The function of the beaver's tail splashing is signalizing danger and got its content due to evolutionary selection processes, being advantageous for survival for those beavers, which were able to interpret the signal as danger. The naturally evolved proper function in this example is the correlation of the event of tail splashing with the occurrence of danger. This does not entail that *every* occurrence of beaver tail splashing has always indicated and will always indicate danger, for beavers being quite shy animals will easily splash in situations without a proper threat. Rather, this correlation means that over time, tail splashing signalized danger more often than being a false alarm, thus saving the lives of many beavers, which in turn was evolutionarily advantageous in terms of reproduction (cf. Millikan 1994)

In Millikan (1995), she describes a special kind of mental representation that can best be described by being descriptive and directive at the same time. She calls them pushmi-pullyu representations and claims that they are more primitive, thus more fundamental for cognition than other mental representations. Mental representations are normally either purely descriptive or purely directive, and require forms of sophisticated cognitive abilities, whereas PPRs are both, but in a more primitive way (cf. Millikan 1995, 186). What Millikan has in mind is that PPRs can be analyzed in terms of having different descriptive and directive content, though this does not imply that the cognitive process using a PPR is doing so:

> Assume further, what is again reasonable, that the effect of the call on the chicks is not filtered through an all-purpose cognitive mechanism that operates by first forming a purely descriptive representation (a belief that there is food over there), then retrieving a relevant directive one (the desire to eat), then performing a practical inference and, finally, acting on the conclusion. *Rather, the call connects directly with action.* (Millikan 1995, 190; my italics)

Millikan claims that PPRs are more primitive, thus more fundamental for cognition than purely descriptive or directive representations, where deployment of the latter is presupposing some practical-inference skills (cf. Millikan 1995, 192). There is good evidence, according to Millikan, that these primitive representations exist. On a neuronal level, mirror neurons have been identified in the motor cortex of monkeys that respond in the same way either to acting on a goal or watching another monkey performing the same task (cf. Rizzolatti et. al. 1988), or on a behavioral level, such as the imitation of facial expressions of newborns (cf. Meltzoff and Moore 1983).

Millikan is able to explain how animals use a primitive representational system for communicating information, and how the meaning of these non-linguistic representational tokens can be accounted for in terms of a biological semantics, involving the notion of proper function. The proper functions in the examples given can best be understood in terms of instinctive behavioral patterns, for which the theory of natural evolutionary selection is a convincing explanation. However, PPRs should not be limited to explaining instinct driven behavior, but also be accountable for spontaneous, dynamic and flexible behavior, based on detection of action opportunities in the environment. Thus, Millikan states:

> The representation of a possibility for action is a directive representation. This is because it actually serves a proper function only if and when it is acted upon. There is no reason to represent what can be done unless this sometimes effects its being done. (Millikan 1995, 191)

Action-possibility representations do have the proper function of action-guidance because, in principle, they can be used by the animal for guiding and initiating actual actions. According to this picture, every representation is a PPR, i.e., has the function of action-guidance that can be used by a consumer mechanism to generate, initiate, control or guide action. This implies that, at least for most animals lacking the capacity of counterfactual reasoning and imagination, that representations representing fundamentally impossible actions could not be processed by this consumer mechanism and could thus not be PPRs, such as representing swimming-opportunity for non-swimmer, or flying opportunity for a non-flyer.

Given that human cognition, already in early stages of infancy, is much more sophisticated than simply locating food sources or detecting possible predators, PPRs for human cognition and their contribution have to be specified separately. Is there is a candidate for a mental representation central to human thought that could be interpreted as a PPR proper, other than beliefs or desires? Beliefs and desires, prominent mental representation types of humans, could simply have evolved from PPRs, being the results of cognitive specification and providing a more differentiated goal and fact representation than PPR could do. However, Millikan rejects this interpretation, assuming that core representations such as PPR, which have been the basic representations enabling further development, are quite likely to have retained their function for cognition (cf. Millikan 1995, 192).

More promising exemplars of PPRs seem to be intentions and the primitive representation of social norms (common norms and role norms), crucially involving desired or required behavior. Thus, intentions clearly have a directive structure, while it is also possible to think of their content as descriptive. Intentions express future goal states and at the same time, they involve the statement that one is about to do what is required for realizing the intended goal state. In that sense, the content of an intention, expressing that something will happen in the future, can be used simultaneously as a description of how the future world will be at a certain time.

Millikan describes intention-PPRs as being different from those previously discussed:

> Rather than functioning as do, say perceptual PPRs, which map variations in the organism's world directly into (possible) actions, it [the intention; T.S.] maps variation in goals directly onto the represented future. It also differs in that the contents of the directive and descriptive aspects of the representations are not different but coincide. (Millikan 1995, 193)

Whereas the perceptual PPRs have two different contents combined in one structure, the intentions as PPRs have two different functions – they can either direct or guide actions to realize a goal state, or they can anticipate the state a world will or might be in at a certain time.

PPRs also occur in representing social norms and roles, although this is not the necessarily the mode of doing so:

> I suggest not that this is the only way humans can cognize these norms and roles, but that it may be the primary functional way, and that this way of thinking may serve as an original and primary social adhesive. (Millikan 1995, 193)

Accordingly, Millikan posits a mechanism that enables humans to understand social norms without distinguishing the directive and descriptive aspects, but integrates both at the same time: most social norms and roles (queuing in lines, being quiet at concerts, obeying teachers, raising a hand when wanting to speak, etc.) implicate imperatives for behavior, while equally being descriptive in that they inform you about the standards and conventions of conduct a society might have. Understanding norms thus can be explicated in terms of entertaining thoughts, which are themselves PPRs:

> It [the mechanism for understanding social norms; T.S] is the capacity and disposition to understand social norms in a way that is undifferentiated between descriptive and directive. What one does […], what a woman does, what a teacher does, how one behaves when one is married or when one is chair of the meeting, these are grasped via thoughts, PPRs, that simultaneously describe and prescribe. (Millikan 1995, 194)

The PPRs that Millikan purports are much more abstract and sophisticated than the primitive perceptual ones. Social norm PPRs describe and prescribe complex social and context dependent behavior, whereas perceptual PPRs transduce features of the environment into possible movements. However, coordinating social behavior is of equal importance to modern humans as safely and purposefully navigating one's (natural, ecological) environment, PPRs for our social environment presumably have the same developmental basis as the perceptual PPRs. Social norm representation can thus be simply interpreted as a more complex way of coordinating one's behavior. The social environment demands for more abstracted representations, nevertheless, these are still part of the developmental continuum of representations.

The final step in showing that PPRs are prevalent in human cognition consists in analyzing language, assuming that if PPRs occur in thought, they are also likely to show up in ordinary language. Thus, certain linguistic utterances can be understood as causing, or evoking, an underlying PPRs. Most declarative sentences have the dual structure of PPRs: 'We don't eat peas with our finger', 'we only cross when the traffic lights are green', etc. These sentences, which could be used in instructing children, have both a descriptive element, which at the same time implies consequences for behavior: The information that generally, people only cross the streets at green light, implies that you are also supposed to do so exactly, whereas a purely descriptive sentence such as 'swans live in lifelong monogamous relationships', does not imply anything for one's immediate behavior. Millikan argues that these examples, and others, such as strict orders, have the

> function [...] to impart an *intention* to a hearer and to impart it directly, without mediation through any decision-making process, for example, without involving first a desire and a practical inference [...], undifferentiated between directive and descriptive, serving to impart PPRs. (Millikan 1995, 194f).

Although Millikan is not specific on this point, it can be assumed with some certainty that the declarative sentences are used *and* understandable only because there is a general mechanism for producing and consuming

PPRs, whose (proper) function it is to guide and coordinate (social) behavior.

A problem for Millikan's account of PPRs in human cognition is that it is quite demanding regarding cognitive abilities and complexity of representational structure. All the examples of intentions and social norms are presupposing conceptual knowledge. On the other hand, Millikan describes the simple, perceptual PPRs, and the cognitive systems generating and processing them, as being crucially involved in understanding of complex PPRs. This is reasonable to assume from a developmental perspective, as important mechanisms with a vital function would normally not simply disappear. However, if this is the case, then an account of transformation or communication is required, that explains how the complex representations arise out of the basic ones and how the conceptual representations are affiliated with the low-level ones. Thus, if some social contexts directly cause intentions, which in turn directly impart PPRs, constituting the immediate grasp of these social situations, Millikan has to explain if the PPRs involved are cognitively of higher order, or if they actually *trigger* or *embed* primitive PPRs.

A solution to this underspecification is, first, to identify primitive PPRs with their neuronal implementations basis, and second, arguing that higher order cognitive abilities are grounded in these neuronal mechanisms. An account of this will be presented in chapter 8 & 9.

Another problem for Millikan's account, and also for consumer oriented accounts in general, might be that an explanation for a specific type of representation is given by postulating a mechanism that is able to exploit these representations. Accordingly, a PPR is whatever can be exploited by an action-guidance mechanism. This is could be interpreted as simply shifting the burden on explaining what a specific consumer system consists of. To avoid this problem, an attempt will be made in chapter 8 & 9 in defining what specifies an action-related representation, by claiming that the representation actually contains motor elements. Thus, representations that are already in a movement format can be exploited and used by cognitive mechanisms, which have the function of controlling and generating behavior (see also ch. 4.2)

5.2 The Guidance Theory of Representation

According to Anderson and Rosenberg (2008), the problem of representational content, or the function of representations, can be tackled with the guidance theory of representations. Their argument for the existence and role of action-guiding representations is of the following structure: There exist a mechanism in organisms that show decoupling of sensory stimulus input and behavioral output generation. This mechanism can best be described as generating and using representations. These representations have the main function of generating behavioral output, hence they are action-guiding (in the sense described below). As these action-guiding representations can be found in cognitive systems that show low degrees of cognitive complexity and only a limited behavioral repertoire, action-guiding representations can be rightly assumed to be a basic cognitive phenomenon that functions as foundation for more sophisticated representational development.

At its core, the guidance theory of representations states that the primary function[19] of representations is to provide guidance for actions – thus, Anderson and Rosenberg focus on what representations *do* instead of asking what they *are:*

> On the guidance theory *R* is about *E* just in case *R* is standardly used by an agent to guide its actions with respect to *E*. (Anderson & Rosenberg, 2008, 57).

Following from this, a cognitive state is a representation if it provides guidance for an agent for executing an action involving the represented environmental object or circumstances. Central to their theory is the assumption of the existence of a 'guidance control system', which makes use of representations that are the consequence of the registration of environmental stimuli (cf. Anderson & Rosenberg 2008, 66). The guidance

19 „What is new about the guidance theory is not that is naturalistic, functionalist, and consumer-oriented, but rather that it insists that the *fundamental ground* of representational content is action guidance." (Anderson and Rosenberg, 2008, 57; my italics)

control system is an evolutionary development of simple behavior-guiding mechanisms. Anderson and Rosenberg illustrate this difference with the example of the slime mold Dictyostelium discoideum, which, in its slug state will always move towards a light source, due to light-sensitive mechanisms in the cells of the mold. This inbuilt mechanism and the resulting internal states of the slime mold are interpreted by Anderson and Rosenberg "to be a prototypical case of the evolutionary pre-conditions that allowed for the emergence of representation-driven behavior" (Anderson & Rosenberg 2008, 61). The internal state of the slime mold drives the behavior of the mold which is evolutionarily advantageous. However, the internal states of the slime mold, being first-order influences on behavior, are not yet to be counted as representational, although they guide behavior. What is missing is a substantial decoupling of stimulus and behavior, which is not given, as the stimulus directly drives the behavior. The slime mold's behavior is important for the discussion of action guiding representations though, as "distinct non-representational but action-guiding bodily states, like the slug's, by being categorized and consumed by a cognitive engine and exploited for self-directed behavioral control, can give rise to cognitively significant representational states" (Anderson & Rosenberg, 2008, 61). Hence, what is missing in order to ascribe action guiding representation in slime mold example is a consuming cognitive mechanism, which categorizes and exploits the relevant bodily states for self-directed behavior.

The example of prey capture in frogs demonstrates how this further cognitive development could look like and therefore establish a case of a minimal representation driven behavioral control/guidance system (cf. Anderson and Rosenberg 2008, 61f). Whenever a small dark, moving dot is entering the visual field of a frog, the frog will turn its body toward the stimulus and snap at it with its gluey tongue. The crucial point here is that the stimulus causes a bodily state (change in retinal ganglion cell firing) which by itself does not trigger or elicit the behavior (frog turning its body to the stimulus in order to snap at it), but is registered by another mechanism (cells in the optic tectum) which then causes the frog to move. The registration is an inner state, which, according to Anderson and Rosenberg, is the consumer mechanism that drives the behavior and thus can count as representational:

> Rather, the stimulations of the retina generated by the small, dark, moving object are registered by, and taken up into, a cognitive system that can consume the registration by exploiting its capacity to guide the frog's behavior in a sophisticated and coordinated way, in context with other registrations. In the slime mold slug there is no such intermediate registration of bodily changes in an integrated control system. This difference is critical enough to introduce the notion of a *potential decoupling* of stimulus and response. (Anderson and Rosenberg 2008, 62)

The decoupling is realized in two ways: different stimuli (thus, not only flies) can trigger the relevant behavior, and the same stimulus can lead to different behavior. The latter becomes evident in the case of a frog, whose optic tectum was removed unilaterally, resulting in the frog to be blind on this side of its visual field, at first. Over time, the optical nerve grew back and attached to the remaining optic tectum on the other side of the frog's brain. From then on, the frog was able to see again in the formerly blind visual region, but processed the stimuli as if they were on the other side of its body, as if it were a mirror image. Accordingly, the frog would jump and snap at the air on the wrong side instead of the actual stimulus position (cf. Ingle 1973). Thus, the same original stimulus gave rise to the reverse behavior, implying that behavioral outcome depends on the action-guiding system to which the stimulus-registration is forwarded (cf. Anderson and Rosenberg 2008, 64). It can also lead to entirely different sets of movements that nevertheless are of the type 'prey capturing': suction-feeding, tongue-snapping, etc. Anderson and Rosenberg are specific about the frog not using the representation of the prey for picturing the world, but only (and, sufficiently) for movement control and guidance. It is not necessary for an organism that the representations it uses for action-guidance represent states or facts of the world as such, as it is sufficient, from an evolutionary perspective that the organism behaves and acts adequately to some stimuli *as if* these stimuli were containing the right kind of information about the world, but it is not really important whether this is the case. In the frog's case, the general function of the representation consuming system is enabling the frog to catch flies and therefore its behavior needs to be adequately guided by the representations and the cognitive consumer system. According to Anderson and Rosenberg, it does not matter for the frog if the stimulus actually has a certain property (being an actual fly as opposed to being a black piece of paper) for its cognitive mechanisms to work and exploit the stimulus properly:

> Although the development of representation-producing and consuming systems was a giant evolutionary leap, its significance is not best elucidated in terms of information-containing, world-reflecting, or situation-modeling inner states. Functionally, these systems are instead best understood as continuous with the older, more-world-driven behavioral systems they replaced: they are the things that provide guidance to the integrated systems for behavioral control. (Anderson and Rosenberg 2008, 66)

The further evolutionary development that gave rise to increasingly sophisticated cognitive systems finally gave rise to more sophisticated ways of representing the world that exceeded the mere use for directly controlling behavior. The same mechanisms for behavioral control and thus action-guiding representations are foundational and at work in all cognitive agents, to various degrees at the different stages of cognitive sophistication and development. Anderson and Rosenberg resume that:

> neither the primary function of registrations, nor the best way of specifying the representations they eventually came to support, radically changed as a result of any of their further evolutionary development. What we see instead are variations on and sophistications of this basic theme. (Anderson and Rosenberg 2008, 66)

This claim entails that cognition is built on and has been developed on the basis of these action-guiding representations. Moreover, it can be concluded that the first representations in cognitive organisms where indeed action-guiding representations, and all other representational capacities and cognitive skills are grounded in these action-related representations. This is not the claim that all cognitive representations are in fact action-related representations, but that action-related representations play a foundational role across species and individuals.

With this basic theoretical framework in mind, Anderson and Rosenberg go on to develop a formal account of the guidance theory to show its applicability as a proper theory of mental representation that captures the complexity of human cognition. At the core of the formal account are the following two definitions:

> *Definition 11*: A token *T tracks* an entity *E* for a subject *S* if, and only if *T* is standardly used to provide guidance to *S* for taking action with respect to *E*.

On the guidance theory, *representation* is simply tracking in the sense defined above.

> *Definition 12*: A token *T represents* an entity *E* for a subject *S* if, and only if *T* tracks *E* for *S*. (Anderson and Rosenberg 2008, 77)

Representation is spelled out in terms of tracking, and tracking implies that the token is used for guiding an action towards an entity E. The content of a representation is thus defined in terms of a directedness of a mental state towards an entity and the fact that it can be, or is, used for interacting with that entity. A mental token would therefore not be a representation if it were somehow about the entity, but could not be used for guiding actions towards that very entity. Anderson and Rosenberg claim that among the criteria, whether something can be used for guiding action, is the possibility of being decoded by a mechanism that is itself "integrated with a subject's action-determining process" (Anderson and Rosenberg 2008, 77). Accordingly, it can be inferred that the token T must be in the right format to be interpretable by the decoding mechanism that generates the action-outcome. Anderson and Rosenberg are not specific about the format of the representations for action-guiding, but they think of the representations being closely coupled with the decoding mechanism. The decoder and the representation thus form a unit, and the more structured the representation is and the more the representation is coupled with what is represented, the less sophisticated the decoding mechanism has to be and vice versa (cf. Anderson and Rosenberg 2008, 77). In this sense, a frog representing a black dot in its visual field as possible prey only requires a simple mechanism generating motoric output from the representational content. The stimulus already specifies the direction of the frog's behavior. In contrary, representing the presence of prey in terms of fresh tracks in the snow requires much more sophisticated interpretative abilities to generate an adequate action-output.

What is the scope of the theory of guidance? As Anderson and Rosenberg already pointed out, they conceive of action-guiding representations as fundamental and basic representations in cognitive development. Furthermore, they want to show that all kinds of representational content can be explicated in terms of its action-guiding potential and thus providing a defining criterion for something being a representation and representational content in general (cf. Anderson and Rosenberg 2008, 77). An example for an action-guiding representation for humans is the case of a driver stopping in front of red traffic light. The driver's percept is a representation of the state of the traffic light ('red'), precisely because this percept guided her action of stopping the car. A young girl's finger counting in order to solve a math problem ('2+3=?') also counts as representation of the numbers involved, because the fingers are used to guide arithmetic reasoning (cf. Anderson & Rosenberg 2008, 79). What about fictional and abstract entities? Can they be representations according to the guidance theory, in that they provide guidance? Anderson and Rosenberg state that representations of fictional entities are representations, and as such are, in principle, able to provide guidance, however, the action system does not respond to their guidance-abilities as they are marked as being fictional. The same representations, would the entertaining subject treat them as non-fictional entities, could well be used for guidance. Apparently, what Anderson and Rosenberg have in mind is a further "judgment" of the subject's cognitive system that a represented content could be real or fictional, and the latter would "mute" the action system and thus actual action-guidance would not occur.

This shows that all kinds of representations *can* be used (in principle) for action guiding, however, it is unclear whether this also shows that action guidance is a necessary (defining) condition for a mental token being a representation. Accordingly, only mental states or states of the organism that could be used for action-guidance would be representations. Anderson and Rosenberg reply to this problem by stating that representational content is not limited to being only immediately usable, but allow for content that matter for action-guidance eventually. This problem can be further relieved if functions other than action-guidance would be introduced, together with broadening the notion of action, e.g. explicitly allowing for 'mental actions' such as inferential reasoning. The example of finger counting guiding arithmetic reasoning can be understood that this could be an option for Andersen and Rosenberg. A general notion of function, involving mechanisms that exploit representations for different purposes, which are not implying motor-actions any longer, is necessary for accounting the cognitive abilities of humans. This in turn needs more elaboration on the role for the development of higher-order cognition on the grounds of action guiding representations.

Anderson and Rosenberg's theory of guidance defines action-relatedness solely in terms of the function of a mental token – whatever is used or has the potential to be used for guiding an action is a representation, if certain other conditions hold (like stimulus-response decoupling). Their account of action-related representation thus heavily focuses on the individual's cognitive processing of environmental information, which is in strong contrast to the accounts that focus on the aspect of environmental properties such as Gibson's affordances (see ch. 3) or Merleau-Ponty's action space of the present world (see ch. 2). It is not of importance, which properties of the environment exactly give rise to the perception of action possibilities, as it is just the use of some information to guide actions that renders the cognitive processing of that information representational. They claim to be independent of a historical description of representational content as in Millikan's (1995) proper functions, however, it seems that Anderson and Rosenberg also have to rely on a historical element to account for representational content. The only way to explain why a given decoder is able to interpret certain information that do not by itself carry action-guiding content (as Gibson's affordances would), is to have an evolutionary development of a decoding mechanism, along with a typical exposure to certain stimuli from which a typical function arises. The frog's prey representation mechanism can hardly be interpreted without appealing to evolutionary developed function on the basis of being exposed to that kind of stimulus, which at least similar to Millikan's idea of a proper function.

Put differently, Anderson and Rosenberg's theory of guidance define representation in terms of guidance alone and thus they do not need to introduce a notion like proper function or something comparable. They simply state that whatever it is being used for guiding an action here and now is an action-guiding representation, no matter if the same representational content was used for guiding in this way in the past. But Anderson and Rosenberg's account also includes a decoding mechanism, which is crucial for the decoupling of stimulus and behavioral output. The function of the decoding mechanism can be best explained in terms of evolutionary development and is thus very similar to Millikan's proper function, re-entering their account through this backdoor. Instead of seeking to provide an alternative to Millikan's notion of proper function, Anderson and Rosenberg could just accept that something as proper function will have to enter any biologically oriented account of representation, which is not too problematic anyway. They seem to be aware of this line of criticism and briefly address the problem in a footnote, where they admit that the

> two methods of fixing content [Millikan's and their own proposal; T.S.] will sometimes, but not always, give the same result. Note the implication, however, that whereas Millikan advocates a direct and prominent role for evolutionary history in determining content, on the guidance theory evolutionary history exerts only an indirect effect on representational content. It is not evolutionary history per se that determines content, but the function of a representation in guiding action. Since it is in virtue of their role in guiding action that the elements of an organism's cognitive systems are primarily exposed to selection pressures, this seems the proper place to locate the influence of evolutionary history on their structure and content. (Anderson and Rosenberg 2008, 83, footnote 16)

This implies that the mechanisms for using a representation for action guidance are based on evolutionary selection and are thus comparable to Millikan's proper function – it is not even clear, if a significant distinction is possible between the two notions, and it could well be that only the focus is different.

5.3 Action-Oriented Representation

Mandik (2005) present an example of an application of how the inner states of a simple system can guide its actions and are thus an instance of action-guiding representations. Mandik calls these representations 'action-oriented representation', a notion borrowed from Colby (1998), who presents evidence for the existence of action-oriented reference frames in parietal cortex. The different representations are related to the eyes, the head, the body, or are hand or grasping related. They can be described as action-oriented because they play a crucial role in guiding motor action towards objects. Mandik's goal is to establish a representational account of 'active perception', which is based on the idea that perception is an activity, dynamically unfolding in the coupling to one's environment (cf. O'Reagan & Noë 2001). To achieve this, he assumes representations, "that include in their contents commands for certain behaviors" (Mandik 2005, 285). He sees action-oriented representations as advancement over existing theories of active perception, which have a too narrow focus on perceptual output conditions, while neglecting the sensory input for perception, and over traditional approaches to perception, which neglect the role of action in perception and consider perceptual input only. Mandik tries to unite these two camps by introducing a notion of perceptual representations that is considering sensory input as well as integrating representations for action, by describing action-oriented representations as contributing to the representational content of perception, while percepts sometimes can be action-oriented representations by themselves (cf. Mandik 2005, 293). He also refers to the special content of action-oriented representations as "imperative content" (Mandik 2005, 293), claiming that imperative content alone would be enough for something being an action-oriented representation and not require, as Clark (1997) or Millikan (1995, see ch. 5.1) do, indicative and imperative content.

As an example for a system using action-oriented representation, Mandik presents the wheel-driven robot Tanky Jr., which navigates by using a simple scanning sensor, calculating the difference between two states of sensor activation. Tanky scans to a position, records the sensory activation at that point and scans to the next position to record the data. If the sum of the two states of the sensor is negative, it means that Tanky is moving away from the light source and has to turn, if the sum is positive, it moves closer to the light source and will continue. There are two modifications of Tanky: the first version uses touch sensors to stop the scanning process, the second version sends a motor command ('scan right') for a fraction of a second, and then sends the opposite command ('scan left'), and so on. Mandik calls the sent (and recorded) motor command an 'efference copy'. Both versions of Tanky do equally well in completing the task, i.e., Tanky is able in both cases to move in the direction of a light source. However, only the second version of Tanky Jr. uses action-oriented representations, Mandik claims, due to the recording of the motor command against which the sensory input variables are computed:

> [In] the efference copy condition [...] the creature knows the position of the scanning organ by keeping track of what commands were sent to the scanning organ. Thus, in the efference copy condition, the percept is genuinely underdetermined by sensation, since what augments the sensory input from the light sensor is not some additional sensory input from the muscles [as in the feedback condition, T.S.], but instead a record of what the outputs were, that is, a copy of the efferent signal. (Mandik 2005, 292)

The resulting representation is a two-dimensional egocentric spatial representation (of the location of the light source), and, due to crucially involving the efferent signal it can be described as an action-oriented representation:

> Thus, in the single-sensor creatures described earlier, the motor command to scan the sensor to the left is as much an adequate representation that something is happening to the left as is a sensory input caused by something happening to the left. (Mandik 2005, 293)

The more general point Mandik wants to make is that the efference copy itself can already be considered being the content of an action-oriented representation. Having the motor command as content is sufficient, though not necessary for something being an action-oriented representation, thus departing from accounts that specify action-relatedness in terms of input and output conditions. Mandik interprets efference copies as already being action-oriented representations, "since they themselves are representations of actions" (Mandik 2005, 302). The role action-oriented representations play for cognition in general is not elaborated in Mandik's account, he is only arguing for the claim that sometimes, perceptual content is already given in terms of an action-oriented efference copy, especially in those cases where the sensory input is underdetermined. Alas, the scope of Mandik's account remains somewhat vague and unclear: are these action-oriented representations alternative strategies in case some sensory information is lacking, or is it crucial for certain perceptual processes? What is the importance for humans and other developed animals, that have available a range of sensory input channels – do they rely sometimes on action-oriented representations, or always in some ways? Mandik does not give answers to these questions. What Mandik provides is a very basic notion of representation of one's environment in terms of combined information from sensory input and motoric output – a way to make sense of sensory input by relating it to motoric output.

5.4 Summary

Mandik's example is only a contribution to modeling action guidance on a very simple representational level, and has thus the theoretical value is rather limited for the present discussion. Millikan's (1995) and Anderson and Rosenberg's (2008) accounts are more substantial for the present discussion, as they emphasize the importance of action from an evolutionary perspective. Both accounts provide convincing arguments, why the primary function of representation is action guidance and detection of action opportunity. They both stress the foundational and derivative role of action-guiding representation for complex cognitive abilities. This can be seen as complementary to the central claim of this book, namely that action-related representation is fundamental for further cognitive development and the cognitive expression of the interactive subject-world relation. However, the accounts discussed in this chapter mainly approach the issue by stating what representations do, whereas the general account of action-related representation should also be able to clarify, what the representations are, what they consist of and if they are special in terms of representational format. Furthermore, the role of the body in representing possible actions, as determining factor for what can be done is not explicitly addressed by neither Millikan's nor Anderson and Rosenberg's account. A theory that seeks to ground cognitive abilities in action-related representations has to account for the role of the subject's body, as the body is the acting instance and crucial part in the determination of opportunities for action for the subject. Thus, both accounts discussed in this chapter neglect how the subject, on the basis of its physical constitution, contributes to representing possible actions and, moreover, how the mechanisms, that generate and control motor action contribute to cognition. This can overcome by introducing a movement format of representation and the role of the body schema for action and action-possibility representation (see ch. 4).

6 Vision for Action: The Two Visual Systems

An important aspect in the discussion of the role of action-related representation for cognition concerns the neuronal structures subserving action cognition. This implies different neuronal areas with different functions: action preparation, generation and the online control of action happens in motor cortical regions, visual areas of the brain processing action relevant stimuli of the environment as well as perceptual input that results from interaction with one's environment, and countless other processes. A central element in many accounts of action- related cognition is the detection and processing of action-relevant information of the environment. Gibson (1986) famously addressed this aspect by stipulating quasi-objective properties of the environment that already specified action possibilities and simply had to be picked up by subjects. As became clear in chapter 3 of this book, this approach has many problematic implications presents thus no viable explanation. Chapter 5 showed that accounts focusing on a functional description of representation provide an evolutionary justification for the general action-relatedness of representation, but lack a characterization of how action-related information is visually processed.

In this chapter, two prominent accounts are presented, approaching the problem of interactive-feature representation (Milner & Goodale 1995) and the further role for representing action possibilities on different levels of abstraction (Jacob & Jeannerod 2003), both proponents of the 'two visual systems theory'. The two systems theory of visual perception states that there exist two neural pathways that are functionally differentiated. The main function for the two visual systems is processing different features of the environment, with one of the systems essentially encoding environmental features in terms of possible actions, thus "translating" visual stimuli directly into a motoric format. The discussion of the two accounts will provide the basis for the further development of a general account of action-related representation enabling a better understanding of how the neuronal mechanisms contribute to different aspects of action cognition (see ch. 8).

6.1 The Two Visual Pathways Hypothesis

The theory of two visual systems was advocated most prominently by Ungerleider and Mishkin (1982), who present evidence for the claim the two streams of visual processing have different functions and thus play a different role in the processing of visual input. The ventral stream is characterized as contributing to the processing of pattern vision, which broadly consists in object identification and recognition. The dorsal stream, in turn, is of importance for visual spatial processing and identifying the locations of objects in the visual field (cf. Ungerleider & Mishkin 1982, 73f). Although the general hypothesis of the functional differentiation of the ventral and the dorsal pathway have not been questioned since Ungerleider and Mishkin's original proposal, more contemporary interpretations deviate in characterizing the functions. Milner and Goodale present evidence for a different interpretation of the functional roles of the two pathways:

> recent findings from a broad range of studies in both humans and monkeys are more consistent with a distinction not between subdomains of perception, but between perception on the one hand and the guidance of action on the other. (Milner & Goodale 1998, 4)

The findings suggest that the ventral stream is sensitive to specific features of objects and provide information about the characteristics of an object – the source for mental processing information about the environment that will form the basis for knowledge. The dorsal stream, on the other hand, can be considered to provide information that is mainly useful for action-guidance. Information such as shape, distance, orientation of an object is encoded here and provides the subject with information that is used by the various motor-control systems. So the two streams are functionally different as in the ventral stream providing ‘knowledge’ about objects in the environment whereas the dorsal stream is sensitive to information concerning motor guidance towards these objects.

The most prominent example in Milner and Goodale (1995) is the discussion of their studies conducted with DF, who suffers from visual from agnosia. DF suffered from severe bilateral damage to her occipitotemporal visual system (what is conceived to be the neural basis of what is called the ventral pathway), while her occipitoparietal visual system (what is, consequently, the dorsal pathway) was left intact (cf. James et al 2003). The symptoms of visual form agnosia include DF being able to control her actions with respect to objects, while at the same time being unable to describe or recognize these objects verbally. For example, DF was unable to report the orientation of a slot that could be rotated by 360°, but was able to correctly insert a card or her hand into the slot, and video recordings showed that her arm and hand immediately began moving with the right rotation when the movement started. The case of DF provides strong evidence for a functional dissociation of the two visual pathways (and their cortical areas which they feed their information to). The findings support the hypothesis that there is a ‘vision for action’ system, which processes, mostly unconscious, information that is consequently made available for the motor system. This information is used for controlling and generating goal-directed behavior, thus, the ‘vision for action’ system basically provides information for action-guidance. As the dorsal pathway is intact in DF, it can be inferred that one function of the ventral pathway involves supplying the cognitive system with invariable object information that is used for object recognition and identification – thus providing a source for world knowledge, knowledge about invariant features of the (objects in the subjects) environment, constituting the basis for stable object representations.

Tum sum up, besides showing that there is a functional differentiation in visual perception, Milner and Goodale (1995) also claim that there is a cortical equivalent that is the cause of the functional differentiation. Put differently, Milner and Goodale provide evidence for brain regions in the visual cortical areas that encode different kinds of information and that are dissociated, i.e., can function (in parts) independently.[20] The ventral pathway thus is the region of the brain that has the function to encode object information that enables subject to identify objects on the basis of their visual features (e.g. shape, color), whereas the dorsal pathway encodes information about the object's action-related features and makes this information available for motor control – the dorsal pathway encodes information that is used for action-guidance. Of central importance for the purpose of the present enquiry is that this way of encoding information is very intimately tied to motor control and thus provides a meaning for the claim that vision is for action. Moreover, the dual pathway hypothesis provides neurocognitive evidence for the claim that properties of objects can be perceived in terms of possible movements, with the dorsal stream processing visual information that becomes directly related to or translated into patterns of movement. This process occurs without the subject being aware at all – no conscious awareness is necessary for selecting the right movement to act upon an object, if the visual information is available for the dorsal pathway and its related cortical areas. The information processed in the dorsal pathway has only function for action-guiding in present situations and thus has no influence on higher-order cognitive states:

> Only this latter, perceptual, system can provide suitable raw materials for our thought processes to act upon. In contrast, the other is designed to guide actions purely in the 'here and now', and its products are consequently useless for later reference. To put it another way, it is only through knowledge gained via the ventral stream that we can exercise insight, hindsight and foresight about the visual world. The visuomotor system may be able to give us 'blindsight', but in doing so can offer no direct input to our mental life […] (Milner & Goodale 1998, 11)

This implies that more sophisticated cognitive operations, such as forming object concepts cannot be based on purely dorsal information alone, but

20 In addition to the two pathways for vision and action, Gallese (2007) argues for the existence of an interference zone, the ventro-dorsal stream. The ventro-dorsal stream, is supposed to serve as the main interaction zone for the ventral and dorsal streams.

need information from other visual channels to establish stable object representations. A possible way to consider the role of the dorsal pathway for operations other than situative action generation could involve providing action-related information for further processing in other domains. Thus, information from the dorsal stream could become integrated with information form the ventral stream (and other sources) and provide the action-related information for an object concept. The visuomotor representations introduced by Jacob and Jeannerod (2003) could be possible outcomes of such integration processes.

6.2 Two Types of Visuomotor Pragmatic Processing

Jacob and Jeannerod take the results from Milner and Goodale (1995) and other findings on the functional differentiation in visual perception and claim that there are two types of visual processing, involving two different kinds of representations:

> we argue in favor of a version of the dualistic approach to human vision. On our view, one and the same objective stimulus can give rise to a perceptual visual representation – a visual percept for short – and to what we shall call a 'visuomotor representation'. Visuomotor representations, which are visual representations of those visual aspects of a target that are relevant to the action to be performed, result from what we shall call the pragmatic processing of objects. (Jacob & Jeannerod 2003, xiii)

According to Jacob and Jeannerod, the function of the visuomotor system is to provide relevant information to what they call 'the intention box', in analogy what many philosophers have called the 'belief box' (cf. Jacob & Jeannerod 2003, xiv). The belief box is fed by information derived (among other sources) from visual perception, whereas the visuomotor system provides the information for goal directed actions, based on visuomotor representations:

> In a nutshell, we claim that grasping the handle of a cup is an object-oriented action, one of whose causes is the agent's intention to grasp the cup. In the course of the action, the agent's intention draws visual information from then visuomotor component from the visual system. The latter delivers a visuomotor representation of the cup that highlights the visual features of the cup relevant for grasping it. One such feature might be the location of the cup coded in so-called 'egocentric coordinated', i.e., in a frame of reference centered on the agent's body. (Jacob & Jeannerod 2003, xiv)

Based on studies done with visual agnostic patient DF (Milner & Goodale 1995), Jacob and Jeannerod elaborate the structure and content of purely visuomotor representations in the endeavor to understand what "to see with a dorsal pathway [alone]" (Jacob & Jeannerod 2003, 185) means.

First, they introduce the distinction of low-level and high level pragmatic processing (cf. Jacob & Jeannerod 2003, 178). *Low-level pragmatic processing* gives rise to basic visuomotor representations of simple actions that are directly related to body parts, such as the hand and the corresponding simple actions of grasping or turning something. *Higher-level pragmatic processing* of visual objects gives enables the subject to perform more sophisticated actions and more complex manipulations and use of tools. In daily routine of healthy subjects, low-level pragmatic processing hardly occurs all on its own, but normally is always occurring together with higher-level pragmatic processing. One of the core elements of the low-level pragmatic processing of objects is that the location of an object is always encoded in an egocentric frame of reference, which enables the subject to guide actions towards the object. Thus, the visuomotor representations that emerge from low-level pragmatic processing contain information such as how to reach for and grasp an object. To put it the other way round, in order for grasping an object, one must represent its location in an egocentric frame of reference.[21]

The egocentric format of visuomotor representation is applied to the case of DF, who is only able to respond to the size and orientation of an object only terms of grasping or manipulating it. If asked for a verbal report without allowing for interaction with the object, DF fails to give a

[21] For a similar argument, see the discussion of egocentric frames of reference in chapter 4.2.

correct specification of the object's size and orientation. She is not able to transfer her visuomotor representation containing egocentric coordinates into more detached representation of the object. Healthy subject are able to shift from an egocentric frame of reference to an allocentric frame of reference and are thus able to determine the size and orientation of objects in relation to other objects. Relying solely on egocentric encoding of an object's location would only allow subjects to determine the absolute size of an object, which is what D.F. is actually limited to. Low-level pragmatic processing thus yields visuomotor representations that encode the location of objects in an egocentric frame of reference and allow for simple actions such as reaching and grasping – actions that are tied directly to bodily features and involve no further, intermediary actions or tools use to accomplish an action goal. All other more complex interactions are also involving higher-level pragmatic processing and mostly also semantic processing of the object (cf. Jacob & Jeannerod 2003, 190). The distinction between low-level representations that encode information in an egocentric way and higher level representations that allow for an allocentric encoding can already be interpreted as an instance of cognitive abstraction, thus describing abstraction mechanisms already on the level of pragmatic representations (for a detailed account of abstraction, see ch. 9).

Low-level pragmatic processing and the resulting basic visuomotor representation only allow for simple actions such as pointing, grasping and reaching. This is of course only a very limited segment of actions that constitute the range of human actions, thus the low-level visuomotor representations Jacob and Jeannerod describe can only account for a small subset what humans actually do. To explain the more sophisticated aspects of human behavior, such as complex manipulation of objects and tool use, Jacob and Jeannerod introduce the notion of higher-level pragmatic processing. Higher-level pragmatic processing crucially involves retrieval and application of action schemas, i.e., stored representations of movement patterns, based on former experience and learning. This becomes salient in the case of tool use, whose manipulation cannot be reduced to simple movements such as grasping, but involves a complex behavioral repertoire:

> Thus, the manipulation of tools includes a higher level of pragmatic processing of the visual attributes of an object than either pointing or reaching. Grasping is necessary but it is not sufficient for the correct use and skilled manipulation of a tool. It is not sufficient because one cannot use a tool (e.g. a hammer, a pencil, a screw-driver let alone a microscope or a cello) unless one has learned to use it, i.e., unless one can retrieve an internal representation of a recipe (a schema) for the manipulation of the object. (Jacob and Jeannerod 2003, 216)

In higher level pragmatic processing, the parietal lobe seems to be crucially involved, forming a part of what Jacob and Jeannerod call the 'praxic system' (cf. Jacob and Jeannerod 2013, 216). The main insights for higher level pragmatic processing and the representations involved come from studies done with apraxic patients. Patients suffering from apraxia have difficulties or are unable to successfully use tools and other artifacts. Their praxic system is damaged, mostly due to parietal lesions. In healthy subjects, the praxic system enables subjects to perform skillful actions with tools which clearly involve higher level pragmatic processing, such as representing the goal of the action, controlling action execution and recognition of actions of other agents as well as imitation of action. Apraxic patients perform poorly at all these things, e.g. patient GW, suffering from a bilateral parietal atrophy, failed to show correct use in all of 15 common household tools, no matter if she was using both of her hands and was verbally instructed or shown the correct use. GW had independent knowledge of the tools and their proper use though, hence she was able to discriminate the tools according to their function and verbally describe the movements involved in using these tools (cf. Ochipa et al. 1997). Another characteristic impairment in apraxic patients is the inability to pantomime actions, such as cutting bread with a knife without there being any bread nor knife – they mainly produce spatiotemporal coordinator errors (cf. Clark et al. 1994). Jacob and Jeannerod (2003) argue that the praxic system is not restricted to action planning and execution alone, but also crucially involved in action recognition of either real or pantomimed actions of other agents, thus being required for all instances of sophisticated action cognition. From this, they conclude that what is impaired in the patients suffering from apraxia is the retrieval of action schemas (stored representations of action patterns) that are needed for skillfully interacting with one's environment. In some cases, as in the case of GW, a set of representations is still existent and can be triggered by visually presenting the tools, but the relevant stored knowledge cannot be integrated into moto plans anymore. In other cases, such as LL's as described by Sirigu et al. (1995), all access to representations of hand actions was blocked.

These case studies show that the ability to pantomime and execute actions with imaginary tools and objects is impaired in these patients, whereas the general ability for reaching and pointing is still intact. The most important conclusion to be drawn from this is that there is either a limited or lacking access to stored action schemas and representations or that the representations and schemas got lost entirely. Thus, the ability to skillfully interact with objects in an agent's environment makes use of formerly acquired knowledge of the interactions with these objects and necessarily involves representations. Simple, basic action possibilities are detected more or less automatically by the low-level pragmatic processing system, involving only low-level visuomotor representations formed "on the fly", whereas complex tool use is a result from learning and thus forming and storing action representations that can be retrieved on other occasions, in the case of pantomime even in the absence of these objects. Higher level pragmatic processing is thus low-level visuomotor representation plus the retrieval of action schemas.

The idea of idea two levels of pragmatic representations presents a challenge to the Gibsonian notion of direct affordance pickup (Gibson 1986; see ch. 3). Whereas the low-level visuomotor processes can be interpreted as picking up affordances, the higher-level processing that crucially involves stored representations and action schemas contradicts Gibson's idea that no mental processes mediate the affordance perception. Jacob and Jeannerod main argument concerns Gibson missing awareness of the dual structure of visual processing, stating that not all of visual perception is about the detection of affordances, but crucially of other perceptual features. Thus, in order to perceptually identify objects in a given scenario, their spatial relation to each other has to be represented in an allocentric way, such as perceiving that the bottle is left to the cup on the table. Affordances, as action possibilities, have to be represented in an egocentric format, which entails that allocentric representations encode object information that is different from the object's affordances (cf. Jacob & Jeannerod 2003, 180f). Furthermore, from the studies with the apraxic patients, it can be followed that the perceptual information available does not enable them any longer to act on the objects affordances, while some conceptual knowledge is still triggered by the visual input. Thus, "pure" affordance detection is not given in these cases, as only a conceptual processing of action-related features occurs, almost without consequences for immediate interaction, which is what the concept of affordances primarily is supposed to explain. In addition, the general idea that stored action schemas are necessary to successfully interact with complex objects, such as tools implies that at least higher-level affordances are represented in terms of stored representation rather than directly picked up by the perceptual system alone, without any mediating mental representations.

Jacob and Jeannerod also use their distinction between low-level and higher level pragmatic processing to point out the limitations of Milner and Goodale's dual systems theory: "Milner and Goodale's (1995) model unduly restricts the role of the parietal lobes to the performance of crude object-oriented actions" (Jacob & Jeannerod 2003, 248). Pragmatic processing is more complex and the role of the parietal role is crucial for retrieval of action-related representations enabling skillful interaction with complex objects and tool manipulation. The model of Jacob and Jeannerod is especially supporting the idea of a gradual transition from basic action-related representations to more sophisticated, conceptual action-related representations in later stages of development, allowing for learning of complex and sophisticated skills and also accounting for more detached representations - embodying the transition from purely egocentric to allocentric representations of objects and action possibilities.

7 Action Constitutes Thinking: Interactive Constructivism

7.1 The Role of Sensorimotor Processes in the Development of Thinking

Another way to account for the role of action for cognition is to model or explicate representation in terms of interaction or sensorimotor processes. This notion can be traced back to Piaget's (1977) idea about the sensorimotor stage in infant development and the role of sensorimotor processes in the development of thinking. The core ideas are that cognitive representations are the result of a combination of sensorimotor skills, processes or competences that are non-representational in nature. Representations are constructed from the low-level sensorimotor processes, thus Piaget is considered to be one of the first prominent advocates or precursors to cognitive constructivism - a view built upon the central idea that knowledge is actively constructed by subjects based on their existing cognitive structures (cf. von Glasersfeld 1990).

A contemporary account in a similar fashion, though more refined and with a stronger focus on the notion of interaction and its role for the development of representations can be found in the work by philosopher and cognitive robotics researcher Mark Bickhard. Bickhard (1999) claims that the foundation of all representation in representational systems is interaction, entailing that only an interactive system can construct or entertain representations at all. Bickhard thus transcends the Piagetian area of early childhood cognitive development and introduces a general theory of interactive representation that is not restricted to explaining human cognition. What is common to both Piaget and Bickhard is that they search for the foundations of higher order cognitive processes or abilities. While Piaget would speak mainly of development of thinking and knowledge, Bickhard is concerned with describing necessary processes that enable representation to emerge. They both consider (inter-)action to be the central element in the process of generating representational knowledge, providing a non-circular model of how representational cognitive processes emerge without presupposing innate concepts or representational knowledge. After a short overview of Piaget's account of cognitive development and the role action plays therein, I will elaborate on the central aspects of Bickhard's theory of interactive representation. Understanding how Piaget thought of action constituting knowledge will make Bickhard's arguments more accessible.[22] Bickhard's account in turn will be useful for elaborating the idea of a gradual abstraction transition in the general account of action-related representation developed in chapter 9.

Piaget proposes that cognitive development of children takes place on stages, with the first stage, from birth to the acquisition of language, being the sensorimotor stage (cf. Tuckman & Monetti 2010, 51). The sensorimotor stage is best described by interactions of the child with its environment is the main source of cognitive development – other forms of cognitive operation have not yet been cultivated but rather arise out of sensorimotor interaction. The primacy of action defines the subject and its world approach of the sensorimotor stage. Thus, the subject can basically be described a set of interaction skills that are directed at objects, involving feedback from this interaction that is processed and results in transforming and further developing the subject's skills (cf. Piaget 1977, 30). The

22 Bickhard on the relation of his work to Piaget's: "Pragmatism in general, and Piaget in particular, worked within a process framework – a framework of action and interaction – and thereby potentially parry Kim's collapse of genuine emergence. Within this framework, they attempted to model, among other things, the nature of representation. I argue, as did Piaget, that representation emerges naturally in the evolution of interactive biological agents, but with crucial divergences in the specifics of the theories. In the theory proposed, representation emerges as the natural solution to problems of action selection and evaluation. Primitive representation, in worms, perhaps, is concerned with relatively unorganized single actions. More familiar kinds of representation – of manipulable objects, for example – emerge in highly complex organizations of interaction possibilities in ways adumbrated in Piaget's constructivism." (Bickhard 2002, 1)

object and the subject do not exist independently from each other, neither subject nor object can be assumed as simply "given" in the child's world:

> the operations of thought derive from action on objects, and every action on the object starts with an indissociable interaction S ⇆ O between a subject S which acts and object O which reacts. (Piaget 1977, 35)

The first cognitive operations and the first knowledge are thus derived from interactions with objects. Crucially, Piaget claims that in the beginning, there is no real subject-object distinction for the child. In fact, it seems that Piaget argues that the subject-object indissociation is a necessary condition for acquiring knowledge:

> The subject S and the objects O are therefore indissociable, and it is from this indissociable interaction S ⇆ O that action, the source of knowledge, originates. The point of departure of this knowledge, therefore, is neither S nor O but the interaction proper to the action itself. It is from this dialectic interaction ⇆ that the object is bit by bit discovered in its objective properties by a "decentration"[23] which frees knowledge of its subjective illusions. It is from this same interaction ⇆ that the subject, by discovering and conquering the object, organizes his actions into a coherent system that constitutes the operations of his intelligence and thought. (Piaget 1977, 31)

In this central passage, the core elements in Piaget's action based account of knowledge and thought become evident. At the beginning, the child forms an indissociable unit with the (objects in its) environment. Objects are part of the child's interactive system, and only by ever growing experience, a slow, gradual detachment is taking place and the objects become more and more independent entities. This process can occur because the object provides the subject with feedback that changes the course of interaction and alters future interactions. Actions become increasingly organized and structured, and in the course of the process, the object as a

---

23 By „decentration", Piaget means the cognitive development where a child slowly moves away from an initially egocentric world to a world shared with other subjects and objects.

structured and independent entity emerges. Basic causality and other physical properties of objects and the world are thus discovered by interactions with objects and the growing structures of action organization. The child thus constructs the object and its own subjectivity in the course of interacting in a progressively organized way. The subsequent development of cognitive operation is formed on the basis of internalization of recursive and revisable actions. The action of combining or grouping objects together (e.g. according to some visually perceptible similarity) and learning that these units can be disassembled afterwards into the original components again form the basis for the cognitive operation of combination or addition, though clearly developing substantial and sophisticated cognitive operations requires a considerably amount of experience and interaction (cf. Piaget 1977, 33). Cognitive operations, knowledge and thought, all three mutually dependent, are thus constructed on the basis of action – starting from reflex-like movements, which provide the first feedback input, to increasingly structured and organized action patterns that establish the subject – object dissociation and define objects and the knowledge of them in terms of organized action structures. Exactly this feature of Piaget's work is also integral to Bickhard's account of interactive representation, which argues for the emergence of representation of the systems interaction with its environment.

7.2 Interactive Representation

Bickhard introduces the notion of *interactive representation* and is rather outspoken about its importance to cognitive science: "*Interactive representation* has claims to be the fundamental form of derivation, from which all others are derivative" (Bickhard 1999, 1) and further, "interactive representation manifests the possibility of being able to account for other prima facie problematic forms of representation, such as objects and numbers, and, therefore, shows a programmatic possibility of being the fundamental form of all representation" (Bickhard, 1999, 13). A central claim of his account is that interactive representation not only accounts for representation as well as misrepresentation, but also for system detectable error, which is a further meta-epistemological criterion all representationalist systems should meet, which can only be successfully achieved by interactive representation. If the other existing representational accounts, such as covariational or functional approaches to representation, are tested on this criterion, they are not only struggling when it comes to misrepresentation (a hallmark of all representationalist accounts), but even more with system detectable error.

System detectable error is a capacity of the representing system to detect that it is actually using or deploying an erroneous representation and is consequently able to learn to apply a better/more appropriate one – relative to its goals. Bickhard picks out two accounts of representation, the covariational and the functional, to demonstrate how they are incapable to provide a convincing strategy to explain misrepresentation as detectable by the system itself. The covariational approach, as proposed, e.g., by Dretske (1981) or Fodor (1990) claim that representational states represent in terms of informational covariance or correspondence of the representing state and what is to be represented. Hence, states representing cows normally covary with the presence of cows – it is cows that cause COW representations, or more broadly, it is the occurrence of cows in the world that the COW representation corresponds to rather than the occurrence of e.g. dogs. These approaches, according to Bickhard, have the general problem of explaining "representational error *at all,* setting aside any issues of the *system detectability* of representational error" (Bickhard 1999, 2).

The problem stems from the fact that in this notion of representation either the covariation or correspondence relation holds, and the representation is adequate, or the relation does not hold, and the representation is about something else or non-existent. A way out for the advocates of the informational covariation was to introduce asymmetric dependence: "The core intuition here is that the possibility of mistaken representations is in some sense *dependent* on the possibility of correct representation; they are parasitic" (Bickhard 1999, 2). The question, whether Bickhard rightly attributes to the "standard representational accounts", such as, e.g., Dretske's or Fodor's, the failure to account for misrepresentation is difficult to address, and a lot has been said on behalf of both the critics of representationalism as well as their defendants, and clearly advocates of representationalism as Dretske or Fodor would be able to offer substantial replies to such kind of criticism as brought forth by Bickhard. As this is not the central topic of this thesis, I'll skip this discussion and focus on Bickhard's proposal: thus, interactive representation (as a kind of action-related representation) can *also* account for misrepresentation and therefore meeting the requirements of a full-fledged account of representation. Furthermore, it is an account of representation that is grounded in interaction skills on the basis of simple movements and feedback processes, involving sensorimotor representations. With such an account, representation and misrepresentation can be explained in due consideration of the developmental origins of representation and thus provide an account that is readily grounded in action.

To start with, Bickhard describes the organization of an interactive system as a system that generally has "some way of indicating the *possibilities* of various interactions that is distinct from *engagement* in those interactions" (Bickhard 1999, 4). Moreover, the system must be able to choose which interaction possibilities, of which there are countless, it will engage in, and does so via anticipation of interaction results – the system thus needs to have "indications of interaction *potentialities,* [and] have indications of anticipated or anticipatable interaction *outcomes*" (Bickhard 1999, 4). Both forms of indications should be the basis for representation while taking care that none of the two forms of indication is presupposing representation itself or be realized in terms of other representations, to avoid circularity problems. How can these indications be specified without introducing representation? What Bickhard has in mind is a procedural theory of cognitive representation that is built on dynamical interaction processes that have the function of keeping the system in a certain "far from equilibrium state" (Bickhard 2002, 8). The focus in Bickhard's account is on internal resources rather than the environmental input, accordingly his notion of representation does not primarily involve representation of facts of the environment of cognitive systems. Instead, interactive representation is oriented towards functionally adequate behavior and interaction with the environment, while being of minor importance what the actual environment is like. In case the environment is different than assumed, the response will not correspond to the predicted outcome that

that would make the interaction selection functionally adequate. In case the environment correspond to the prediction, the interaction selection was functionally adequate. These criteria guarantee the success of interactive representation, while any sense of being true or real is unimportant.

The model of interactive representation has three central components from which representation emerges: indicated interaction potentiality, indicated interaction outcome and detecting environmental change/environmental features. Representational systems are recursive self-maintenance systems, which are able to register changes in an environment, and on the basis of that registration the system indicates interaction possibilities that imply predictions about possible outcomes. Let' have a closer look how these components work together.

First of all, Bickhard restricts the indication of interaction possibility to the indication of interaction *types* instead of interaction *tokens.* Interaction types in turn

> "are easily specified by the functional or control structure organizations that would engage in those interactions, should the system select them. Interaction types, then, can be indicated by indicating subsystem organizations, like subroutines or servomechanisms." (Bickhard 1999, 5)

The indication of system components happens via pointers:

> A collection of pointers in a privileged location that point to subsystems will suffice to indicate the interactions that would be engaged in by those subsystems as currently available. (Bickhard 1999, 5)

Apparently, what Bickhard has in mind is defining interaction types, i.e., possible movements the system is capable of executing, in terms of the subsystems that would generate, initiate and control these movements. The motor control system with its feedback loops and efferent copy is an example for such a system. Thus, a representational system has a variety of possible interactions at disposal. Every interaction possibility goes along with a prediction of a possible outcome, i.e., a future state where some sort of goal state is realized. The goal states do not need to be representational (which would introduce circularity), but could be explicated in terms of

> internal set points for a servomechanism process that selects one space of internal processes if the set point is not met and a different space of internal processes if the set point is met. (Bickhard 2002, 13)

The system is also able to detect inappropriate interaction: when an interaction type does not meet the predicted outcome state, the system will not be in the desired state and will simply register the failure of its action and go over into another state indicating a different interaction. From this, the system is able to self-detect inappropriate actions in certain environmental situations and is thus able to learn and adjust. Different action potentials can be indicated for more complex systems in the same situation with varying predicted outcomes and by exercising some of the interaction possibilities, the system will learn which one leads to satisfying the desired goal state. Representation in Bickhard's account is therefore a synergy of indication interactions (selecting from a movement repertoire), predicting interaction outcomes (predicting the state the system will be in, in terms of sensorimotor feedback contingencies) and detecting environmental circumstances. As an application for this model of representation, Bickhard gives an example how simple object representation could develop on these grounds:

> Consider, for example, a toy block. A child can do many things with it, from visual scans to manipulations to chewing to throwing, and so on. If any of these are possible, then all are possible, perhaps with intermediate interactions, such as a manipulation to bring a particular visual scan back into view. Furthermore, the entire web of interactive potentialities that the block affords remain invariant under a large class of physical interactions, such as hiding, leaving in the toy box, walking out of the room, and so on — though it is not invariant under such processes as burning or crushing. From an epistemological point of view, *this is a small manipulable object.* (Bickhard 2002, 13; my italics)

An object is constituted by a special organization of a web of interaction indications – the way these interaction indications are branched and determine each other, e.g. through intermediate interactions. As basically every object offers a set of interaction possibilities, or to speak with Gibson, has multiple affordances, each object is represented in terms of all the possible interactions a system can indicate plus the way the potential interactions for a system are organized by this system. An object is thus nothing more than a set of structured interaction indications. Bickhard's idea appears similar to unsupervised learning in neural networks, where operations such as clustering or object and pattern recognition are based on self-organization, with the purpose to detect structural properties of the input domain and consequently adapting network's internal structure to these properties (cf. Ultsch 1993).

Another important aspect in Bickhard's account is that he conceives of the indication process as purely internal: pointers, i.e., system processes point to subsystems organizing and controlling movement – the subsystems being motor routines or something similar. The indication of action-outcomes also has to be internal, if it was the anticipation of external outcomes, these would have to be represented and then Bickhard's account would become circular. Indication of interaction outcomes are thus indications of internal states that the system will be in after undertaking the interaction. Representation as such is an internal process: the indication and anticipations of interactions, combined with the presuppositions they include about the interaction outcome determine representational content, the presuppositions which can either be satisfied or not be satisfied by external conditions:

> If content is determined by the representation itself, however, *independent* of the represented, as it is by interaction anticipations, then there is no problem with the representation of non-existents. That is, the content is internally determined by the representational anticipations, and there is no need that anything exists to satisfy the presuppositions involved in order for those presupposed conditions to be presupposed. (Bickhard 2003, 5)

This is the reason why it is possible for the system itself to detect when it is in error: unlike (standard) representational accounts which stress the relation of the mental representation and the represented, Bickhard focuses only on the internal interaction anticipation, which consists of an implicit prediction about the environment, however never explicitly represents the environment. This prediction can be simply false and the system can detect its falsehood by actually undertaking the interaction and detect that the indicated outcome is different from the one that actually obtains:

> Simultaneously, such organizations of indications constitute even more sophisticated representations. An indication of potentiality is still an implicit predication and is capable of being false. In this organization, there is also the possibility that the system can itself discover such falsity, should the indication be undertaken. In particular, if the *actual* outcome is not among those indicated, then the indications were false. Such error can be useful for further selections of interactions or for invoking and guiding learning processes [...]. At this point, we not only have representational error, we have system detectable representational error. (Bickhard, 2002, 12)

The idea of falsehood detection on the basis of prediction and outcome can be found in a similar way in Barsalou (1999), where he describes the abstract operation of 'negation' with a mismatch in anticipated representations and the actual input from perceptual representations.

Bickhard is offering an account of action-related representation that only implicitly represent features of the environment as features. Rather, the features of the environment are only part of the triggering or detection of change, and are implicitly represented in the interaction-outcome indication. To stick with one of the few examples Bickhard gives, a frog would represent a black moving dot not *as a fly or as prey*, but would represent the fly only *implicitly* in terms of 1) a detected and selected interaction possibility (e.g. moving the head; flicking out tongue) and 2) the indicated/predicated *internal* interaction outcome (e.g. getting the right sensory feedback from the tongue; having the indicated change in the visual field). If the interaction selection was inappropriate, then the indicated outcome will not obtain and enable the system to detect the inappropriate selection and select a different interaction indication from the set of possible interaction indications – if available.

Bickhard can thus account for a primitive mode of representing features in terms of possible interactions, and with the feature of internal interaction outcome indication, he can also provide an account misrepresentation – or can, in other words, provide an account of representation at all. Without the possibility to be in error, states of a system that are related to its environment can hardly be representations and thus the criterion of system detectable error is of crucial importance for Bickhard's account. But misrepresentation aside, what is most valuable is Bickhard's acknowledgement that representation has to start *somewhere* – i.e., a cognitive system cannot just acquire representations on the fly without having representation already, which always leads to skeptical arguments embracing implicitly or explicitly foundationalism (cf. Allen & Bickhard 2013). By foundationalism, Bickhard refers to positions that are forced to stipulate innate modules, concepts or representations to explain further acquisition of knowledge (cf. Chomsky 1959, Fodor 1975). (Neo-)-empiricist accounts have tried to address and solve the problem of the foundations of knowledge since its rise in the 17th century, however, according to Bickhard, did not provide a satisfactory solution so far.[24]

The interactive approach to representation proposes to explain representation in terms of anticipation rather than in terms of correspondence, as anticipation is supposed to be specifiable in functional terms and thus does not need any representational base (cf. Allen and Bickhard 2013, 127). The interactive approach is thus an explication of one of Piaget's (1977) central claims, the idea that representation emerges from a non-representational base. The non-representational base is comprised of "certain goal oriented motor capabilities and representational knowledge is an emergent product of constructions that use them" (Allen and Bickhard 2013, 126). The further development consists in specifying the goal-oriented motor capabilities in terms of action-outcome anticipation. Anticipations of future states that would obtain if the action would be selected exemplify the two central features of representations: aboutness and truth-value. Anticipations are about possible states of the organism and the world, and

24 Cf. Barsalou (1999), Prinz (2005) for advocating neo-empiricist positions on concept acquisition and Allen and Bickhard (2013) for a critical review of empiricist approaches to providing a foundation for representation.

they can obtain or simply be wrong, hence they feature aboutness/intentionality and truth-value.

To sum up: Systems capable of flexible interaction with their environment (necessarily) use anticipation of action outcomes. The anticipations presuppose environmental conditions that would enable successful interaction. The anticipations itself are no representations yet, representation emerges when the anticipation is used for determining the success of an action and shapes future interaction. Anticipation thus needs the actual motor action to be either validated or falsified, therefore gaining representational content – in a minimal sense.

Unfortunately, Bickhard is never really specific when it comes to spelling out the details what it exactly means to define anticipation functionally. One way of interpreting it is that Bickhard assumes a system is not entirely restricted to anticipating the next (internal) state the system will be in after acting, but also allowing for a minimal prediction about the states of the environment. The frog, flicking its tongue at a black spot is anticipating a state in which the frog as a system maintains functionality, and by doing this, the frog anticipates a feature of the environment too. If the tongue-flicking proved successful in the sense that inner states of the frog signal success (e.g. feedback from the digestive system, blood sugar level etc.) the situation will be stored as one in which tongue flicking is appropriate and representational knowledge in a minimal sense about the frog's environment has been acquired.

A problem with this construal of the representational base is that Bickhard never provides a striking argument why anticipation is not representational already. The claim that representation emerges on this basis is thus underspecified and it seems that Bickhard cannot give any example where representation truly emerges from some non-representational states. One could always interpret the anticipations of environmental conditions as representing the environment in a certain state relative to a certain state of the system, thus it is not clear why this interpretation should be dismissed.

There might be an alternative explanation of how representation could have emerged on the grounds of unspecific movement, thus avoiding the circularity worry with anticipation. By assuming that the first representation naturally emerge with the first movements and the feedback processes resulting from this, it could be explained how an organism slowly starts building representational knowledge about its environment. What has to be presupposed is a functioning sensory system, such as visual and proprioceptive input channels for tracking and recording the self-generated feedback. The goals for goal-oriented action can be given by the situation or be even considered to be hardwired to a certain extent, as an organism needs to ingest energy, and evolutionary selection might have equipped organisms with the first "goals", maybe even in the sense of stored, innate information. This is by no means a complete account of representation development, but rather a speculative proposal of which factors could or should be considered for the foundations of cognitive development. As this question cannot be dealt with in a satisfying way due to complexity of the problem, a way to move on is simply bypass the question of the first emergence of representation and to focus on the further development – from simple detections of interaction possibilities to object representations consisting of bundled, structured and organized interaction possibilities, resulting from motor output and feedback processes. This way, Bickhard's strong focus on interaction for representation can be made applicable for explaining the development of higher order representation (and even conceptual knowledge) on the basis of interaction and the subsystems involved in successful goal-oriented interaction-generation and the evaluation thereof.

Another potential weakness of Bickhard's interactive account is that he almost entirely leaves out perception and perceptual representations. He only accounts for sensory input in terms of 'interaction potentiality detection', which clearly involves informational input from the environment and the very features that possibly give rise to interaction possibilities at all. Thus, his account provides an interesting starting point for the development of very basic cognitive representations, but at some point the complex perceptual representation of all sorts of environmental features that are at least not obviously action-related (e.g. shape, color) should to be considered to a greater extend. This could be done by reference to the two visual systems literature (see chapter 6), which provides strong evidence for a vision for action channel and thus highlights the importance of visual input for interaction selection. According the two visual systems theory, the dorsal pathway, terminating in the parietal lobe, has the function of visually guiding behavior. This implies that some action possibilities are automatically processed and made available for the motor system by the dorsal pathway, and emphasizes that cognitive systems perceive environmental features according to possible actions. This is no counter evidence to Bickhard's account, but should be seen as complementary theory with a stronger focus in the input conditions, whereas Bickhard puts more emphasis on the output conditions.

Bickhard and Piaget's interactive approaches are addressing the question of the development of general representational abilities of cognitive systems, which is an important contribution to the debate about action-related cognition. Many theories of action-related cognition are mainly focused on the (neuro-) cognitive processes underlying action-relevant information processing, such as described in chapter 6. Other accounts such as those concerned with the evolutionary function of representation in general (cf. chapter 5) focus on representational content in terms of action goal realization. Interactive representation can be a valuable addition by providing a theory of representation generation. In the next chapter, these (and other) aspects of the different accounts of action-related representation will be taken as foundation for developing a general account of action-related representation. The aim is to come up with a notion of action-related representation that brings together the various aspects of action-related representation discussed so far and by doing so, providing a more refined version of action-related representation that is able to explain the behavior of a wide range of animals while at the same time accounting for the role of action for the individual cognitive development of subjects.

8 A General Account of Action-Related Representation

In this chapter, an account of action-related representation will be developed, based on the discussion of the various accounts of action-related cognition in the previous chapters. The general account of action-related representation is a substrate as well as a refinement of the previous accounts. This will be the basis for the discussion of abstraction processes at work in action-related cognition that will yield a new understanding of how abstract cognitive development can be explained in the light of theories of grounded and embodied cognition (see ch. 1 for an overview). But first of all, it has to be clarified what action related cognition is based on.

The main claim is that action-related representation at different levels of sophistication and complexity enables action-related cognitive processes, such as planning and guiding goal-directed action, perception of action possibilities as well as the perception and understanding of other agents' actions. Action-related representation has two aspects: the 'internal' processes of action generation, action planning and motor intention, and the 'external' factors, such as features and properties of objects and situations. Most accounts discussed in the previous chapters focused on one aspect over the other: Gibsonian affordance perception (1986) focuses on the external relations between perceiving subject and the objective features of the environment, whereas the neurocognitive accounts of Milner and Goodale (1995) or Jacob and Jeannerod (2003) focus on the internal neurological processes involved in visual perception of action-related features. Piaget (1977) and Bickhard (1999) are mainly concerned with the output side of action, and less concerned with describing the features of the world that actually enable subjects to interact with successfully, together with a strong focus on cognitive development of representations in general on the basis of movements and interaction. Accounts with a focus in the function of representation for action guidance, as discussed in chapter 5 have the goal of explaining representational content with action guidance, providing an alternative to informational accounts of representation (cf. Fodor 1975, Dretske 1981; 1986). Chapter 4 focused on explaining the role of self-related and egocentric representation for representing action possibilities for subjects in general, while also arguing that the information about the body of the subject crucially has to have an influence on the way action possibilities are represented by the subjects. To bring these different aspects together, I will make the following proposition for a general account of action-related representation:

If a subject represents action-related features of an object, it entertains or deploys an action-related representation. A feature of an object is an action-related feature (for the representing subject) if the feature of the object is represented in the format of a possible movement. In analogy, if the object's feature would be represented in a visual format, it would be a visual representation of the object's features: the same logic applies to the other sense modalities. Representations are most often multi-modal. Normally, representing an object in terms of the goal-directed movements it allows for, also implies a conscious visual representation of the object's task irrelevant features such as its color or higher-level features such as its price. The dissociations described by Milner and Goodale (1995), Jacob and Jeannerod (2003), and possible others provide evidence though that action-related processing can be detached from purely visual representation of an object's features. A more formal account of this definition would look like this:

> Representation R is an action-related representation if:
> R represents feature F (of an object) as possible movement M of an agent A (in accordance to bodily aspects B of A).

According to this definition, representing the handle of a cup as the endpoint of a reaching and grasping movement (relative to A's arm length and hand span) is an action-related representation of the agent entertaining this representation. This definition captures only the most general aspects of an action-related representation – it is simply the basic structure, which can be and will be at times enriched to varying degrees. The notion of action-related representation as developed here, allows for a hierarchical organization of action-related representations from simple actions to more complex ones, and exemplifies cognitive abstraction mechanisms on the level of action-related representation. This is made manifest in the transition from purely egocentric, implicitly representing possible actions for the agent, to allocentric, explicit action representation for whole classes of agents.

Further essential aspects of action-related representation, which have not been explicitly mentioned in the above definition, are as follows: goal representation; egocentricity; and the prerequisites of action-related representation. The latter involves concepts possession, intentions and other contextual features as possible conditions. These are now to be described in detail.

8.1 Features of Action-Related Representations

8.1.1 Goal Representation

Goal representation is necessary for something to count as an action – among the minimal defining criteria for actions in general, is goal-directed behavior. The action goal has to be represented somehow, otherwise the definition would face the same shortcomings as did Gibson's (1986) affordances, in being unable to explain *why* a certain feature in an animal's environment becomes relevant or motivational at all (see ch. 3). This brings in the notion of motivation or intention for action, which is closely connected to the action goal – normally, that what is intended is a certain goal-state to obtain. One can of course represent a possible goal-state without having any motivation or intention to act accordingly, but the reverse seems rather impossible: entertaining an intention to act without intending a possible goal-state to obtain is unlikely and at least restricted to special cases.[25]

[25] One could imagine someone intending to do A without knowing what the possible consequences would be – as in someone taking an unknown drug without having

One way to account for goal representation is to include it in the generation of the possible movement. A grasping movement towards an object has the hand-object contact as endpoint, so the goal-state of a successful grasping movement is the actual object-contact. The movement cannot be generated if the starting point and endpoint are not given, or impossible to specify. Representing an object in an action-related way as graspable actually means representing the object in terms of a possible endpoint of an action. Here, the role bodily aspects play in representing something as act-on-able becomes obvious: a possible grasping movement implies the arm length and hand span of the agent – grasping normally means reaching out an arm length and forming a grip around an object or some of its parts. Thus, an agent, in the simplest case, will represent objects as graspable in accordance to the agent's bodily aspects. The goal representation could be realized already on the level of basic neuronal mechanisms. There is evidence from primate studies by Hoshi and Tanji (2000), who show that in planning and preparing a motor task, information about the target and the relevant body part must be integrated prior to the generation of a motor command that initiates the appropriate limb movement. Among the possible involved structures of the brain for integrating these two different kinds information, the premotor area in the cerebral cortex of primates could be identified. The lateral sector of the dorsal premotor cortex processes both visual and somatosensory input, involving neurons of this area that gather information about both the action-target and the relevant body parts, such as the arm in a grasping task. These findings suggest that goal selection in accordance to bodily aspects happens even prior to the generation of a motor command. Mechanisms like this are presumably among the instantiation conditions of action-related representations.

any idea what the result state could be, or someone wanting to throw a stone at a window without knowing that this will possibly result in a broken window, but nevertheless wanting to throw the stone. Although, these and other examples are conceivable, they clearly seem to be out of the normal and should thus be treated as exceptions to an otherwise generalizable rule.

8.1.2 Egocentricity

Basic action-related representations are essentially indexical and egocentric. An action-related representation is, on the most basic level, an egocentric representation, and egocentric representations in turn are already represented in an *action-format* – and vice versa (cf. Vosgerau 2009; see ch. 4). This can be illustrated with the fact that representing an object's feature in terms of movements entails that these are the movements of an agent. For the agent, the natural thing to do is to represent the object in terms of her own movements, as representing the movements of other agents is already involving abstraction and should not enter the general description of action-related representation. The action-related representation 'this is graspable' is essentially indexical in that it is about the agent's movements and the agent's perspective only – any other agent would represent a different object or situation with this action-related representation. Hence, basic action-related representations are always egocentric and only valid for the representing agent. The agent (her skills, her bodily constitution) is thus always, at least implicitly, represented in every action-related representation. It would not make sense to assume that the connection from the action to the agent has to be made explicit, by say, inference. This extra step is unnecessary for explaining simple interactions with the world, such as in reaching, grasping, pointing or ducking to avoid objects.

8.1.3 Action-Related Representation Presupposes no Concepts

Connected to the previous criteria of egocentricity and indexicality is the criterion that action-related representation is itself non-conceptual – at least on the basic level. To form, entertain and act on an action-related representation does not imply mastery of conceptual reasoning skills or complex inferential skills. If this were the case, this would limit action-related representations severely in their explanatory scope – excluding not only all kinds of animals, but also human babies and toddlers. The best way to explain the interactions of cats, chimpanzees and human toddlers with their environments is by appealing to action-related representations.

Representing an object as graspable or reachable does not require at any level being able to represent the object as the object it is, nor does it require representing the type of action – such as explicitly representing a bottle as within reach, in terms of representing the bottle as a bottle that can be reached by my reaching and grasping action. This is not to state this kind of mental process cannot be involved in action cognition and is sometimes the most adequate way of cognizing a situation; it is to state that this kind of higher level cognitive operation is just not needed for representing an object as graspable and successfully reaching for it. Distances are given in action-related terms: from very early on, without knowing what distance as such means, animals have a meaning for reachability, which refers to objects in their reaching distance – none of the components have to be made explicit in representing something as reachable. It follows that no concepts are needed and the action-related representations at work can be as simple as possible and still successfully guide and explain flexible behavior qua being representations.

Furthermore, as research on tool use in animals shows, animals of various species are able to use different tools for goal-directed behavior. Chimpanzees use sticks to successfully push peanuts out of tubes while preventing the peanut from falling into a trap (cf. Povinelli 2000, 110). New Caledonian crows use hooks for food foraging and, even more astonishing, create hooks out of bendable material to use as a feeding tool (cf. Weir & Kacelnik 2006; Hansell, 2007). The animals are thus not only able to represent the affordances of the tools, but are also able to invent tools according to their goals and purposes. This kind of flexible and creative behavior can be explained by referring to action-related representation, which includes the integration of bodily aspects, skills and stored knowledge that can be applied to novel situations. However, leaving open the possibility for the possession of structured representations, which itself might be pre-conceptual, most animals are not supposed to possess a complex repertoire of concepts. Thus, the action-related representations that explain the behavior of these animals do not by themselves presuppose concepts, but can rather be understood as the starting point of at least some processes of conceptual development.

Moreover, from behavioral and neurological evidence it can be concluded that the processing of basic action-related features is, to some extent, automatic and not (necessarily) the product of conscious mental processes (cf. Milner & Goodale 1995, Ellis & Tucker 2000, Jacob & Jeannerod, 2003).

Thus, action-related representation is a fundamental component in organizing and guiding the behavior organisms, whose behavior shows minimal flexibility at least. Action-related representations are basic and simple, and can therefore be implemented by primitive mechanisms in many animals. Of course, higher-level action cognition, such as complex tool use and manipulation or understanding other agents, needs more complex structured representations, but these are still in line with, and originate from, basic action-related representation. With the development of conceptual thinking, the representational power increases drastically and increasingly sophisticated actions can be executed – as well as many other sophisticated cognitive skills.

8.1.4 Intention for Action

Philosophical action explanation will normally include some intentional aspects of the agent, such as the belief-desire model of action explanation (cf. Davidson 1967). The intentions of the agent are thus an essential part of the explanation for the action outcome. Action-related representations are independent from intentions in the sense that they are merely representing a possibility. The intention to act enters the stage when the actual action generated on the basis of the action-related representation is executed. The intention can be part of the general action context for an agent, sometimes it is sufficient to speak of motivational and situational features that explain why an action occurred. An animal suddenly being confronted with a predator will immediately have to react and therefore search its environment for escape routes. The animal's perceived action opportunities are: 'allows for hiding', 'is a passable way', 'is an enclosed space' and the like. An animal representing its environment in this way is using action-related representation, guiding its action in accordance to its motivational situation.

Intentions are typically conceived of being mental states involving propositional content, which involves possession of concepts. Thus, rather than attributing intentions to animals, the broader term 'motivation' should be used instead. Accordingly, action-related representations are formed or executed due to a change in the animal's motivational situation. Hence, there is a logical independence of intention and action-related representation. For instance, action-related features of objects are processed even if they are task irrelevant and therefore not part of the agent's intentions to act (cf. Ellis and Tucker 2000). This suggests object features can be processed in an action-related way, without the agent even planning to act on these object features. It remains to be shown if this includes only a small set of basic actions and object features that are linked to bodily properties (such as: 'handle-grasping') or if this phenomenon can also be found at more complex levels of feature processing.

On the other hand, intentions and motivations do crucially drive the selection of actions. Studies by Craighero et al. (1999) could demonstrate that subjects who prepare for a grasping action are faster in detecting a grasping possibility. In their experiments, subjects had to respond to visual stimuli by grasping a handle. Being prepared for a grasping movement enhanced their performance when the stimulus was congruent with the prepared action. Thus, preparing for action enhances the visual selection of the relevant action cues. This supports the idea that intentions to act are cognitive states guiding the selection and detection of action-relevant features.

What exactly leads to intentions and motivational states is not the topic of this chapter, but it can be safely assumed that intentions form based on several sources: biological needs, psychological dispositions, situational features, hormonal changes, newly acquired information or the results of reasoning and conceptual thought. These sources all contribute to changing and influencing the motivational states of organisms. Action-related representation does not directly involve representation of intentions, whereas the reverse is true. An intention for action involves, almost always, an action-related representation. To intend to do *f* includes representing *f* as doable, and thus implies in turn that both the object-features involved in *f*-ing as well as the subject's abilities are also represented in the action-related representation of *doing f*. If a subject intends to hang up a picture, she will generate all sorts of action-related representations on the way, as the action involves many steps. She would probably check the weight of the picture and decide if a nail would be sufficient or if a hole needs to be drilled for anchoring a screw. This might lead her to knock on the wall at the supposed future site of the picture, to conclude from the auditory and tactile feedback if the wall is strong enough and not a light and hollow plaster construction. These are all examples of action-related representation, admittedly on a higher level that the previous grasping and reaching examples. Nevertheless, representing a nail as strong enough to carry the weight of a picture is an action-related representation, as it represents the object in terms of the many possible actions it will allow for: being nailed into the wall, holding the weight of the picture, etc.

This way of analyzing intentions in terms of action-related representation is furthermore able to add some 'grounding' to the standard belief-desire model of action explanation (cf. Davidson 1967). According to a simplified version of this model, the action to open the refrigerator can be explained on the basis of the desire to drink a cold beer and the belief that there is a cold beer in the refrigerator. This answers the question why the subject does *f*, but is completely silent on the question of how the subject actually achieves to *f*. A possible solution could be: By generating and deploying a succession of action related representations, in accordance with the intended goal-state. This might sound overly complicated for the refrigerator-beer example, as this is probably an almost automatic process – subjects open refrigerator doors countless times in their lives. Nevertheless, it crucially involves reaching and grasping actions, which rely on simple action-related representations of the handle and the bottle as graspable. While most of what people do on an average day is automatic, this does not mean that the cognitive processes underlying these automatized actions are not sophisticated. Imagine the case where the subject wants a cold beer, and knows that a cold beer is in the refrigerator, but the refrigerator breaks while attempting to open the refrigerator. Unfortunately, this particular refrigerator is an urban vintage style American refrigerator that actually has a locking mechanism, which opens by pressing down the handle - which is no longer functioning. The subject, still being thirsty after a long day in the office, writing her thesis, needs to find an alternative handle and scans her environment for an adequate utensil, until she finally finds a screwdriver that looks as if it would fit into the mechanism and can be used to open the refrigerator door – and she is successful. This example should illustrate that describing a situation of action on the level of action-related representation is entirely different from describing it merely on the level of the belief-desire model. The standard belief-desire model neglects all dynamic and contextual aspects interaction-situations typically feature. The claim is thus: that which the subjects do in these situations and how they achieve the intended goal states can only be accounted for by referring to action-related representations.

8.2 Empirical Support for Action-Related Representation

In this section, further empirical evidence, will be presented, in addition to what has already been mentioned in the respective sections, to support the idea of general action-related representation as an important aspect of cognition.

8.2.1 Body Schema

As discussed in chapter 4, information stored in the body schema (cf. Gallagher & Cole 1995) is essential for movement in general and especially for successful interaction with the world. Motor plans and commands are generated in accordance with information stored in the body schema, such as position of the subject's limbs at a given time and the body's posture, etc. The body schema is not only operant in action execution, providing information about the current state of the body for adequate movement generation, but also in representing action possibilities. To represent an object as graspable, it must be in reach in relation the subject's position and arm length, yet also match grip size. Information like this is 'projected' onto the world and action possibilities are thus detected – actions are possible for the subject in accordance with properties about the subject's body. The body schema is dynamic and flexible and reflects changes of the subject: growing in height, acquiring skills or disabilities, and similar changes are dynamically updated. This plasticity of the body schema even allows for integrating tools (Carlson et al. 2010) or allows for relatively fast reorganization as demonstrated in the famous rubber-hand illusion experiments (Lewis & Lloyd 2010).

Tool integration in the body schema is of special importance for specifying action-related representation. Studies analyzing neuronal activity in the intraparietal cortex in macaques provide evidence that, after some training, the macaque body schema is updated to include tools, such as those used for reaching, into the body schema (Maravita & Iriki 2004). There is also evidence that the body schema plays an important role in both simple and complex tool use in humans, far beyond that of macaques, even without the necessity of extensive training phases (Berti & Frassinetti 2000; Johnson-Frey 2004; Chaminade et al. 2005).

From this, it follows that the body schema is crucial for tool use because information regarding interaction with tools is immediately or at least rather quickly integrated into the body schema. This is the basis for detecting action possibilities, as the body schema not only stores and provides information that is body part related, but also integrates external objects features into the spatial moto-behavioral representation of the body. The body schema is thus open, flexible and dynamic and allows for the integration of new information, mainly on the basis of proprioceptive feedback. With the body schema active, a subject is not only able to determine immediately if she fits through a gap or can reach and grasp for an object of a certain width, the subject is also able to judge if a stick would be an adequate elongation of her reaching length. Body schema information allows for generating adequate motor commands for goal directed behavior and furthermore enables detection of action-relevant features that are related to bodily properties, skills and prior experience with tools and artefacts. Action-related representation therefore crucially involves reference and deployment of information stored in the body schema, as the body schema represents the body in its three-dimensionality.

8.2.2 Vision for Action

Related to the discussion of the role of the body schema is the function of the dorsal pathway in visually processing action-relevant features. As Milner & Goodale (1995) and Jacob & Jeannerod (2003) have shown, the visual system is functionally and anatomically divided in two sub-systems, processing different object-related information (see chapter 6 for a detailed discussion of both accounts). In defining action-related representation, the evidence presented for 'vision for action' becomes a central element, as it enables a better understanding of how humans process visual, object-related stimuli. Object features that can be related to simple actions are thus visually (and non-consciously) represented by functionally identifiable cognitive processes and are immediately made available for the motor system. According to the notion of action-related representation, visual input is processed in an action-related way, in terms of possible movements. 'Vision for action', and the resulting low-level visuomotor representation (Jacob & Jeannerod 2003) thus provides the basis for the object-feature representation element in action-related representation and shows how the idea of 'representing in an action format' can be spelled out on neurocognitive, functional level. The information processed in the dorsal pathway mainly has the function of generating appropriate motor commands and is thus not available for other cognitive processes, such as conscious processing of the stimuli. It seems fair to conclude from this that a part of vision is concerned with encoding environmental features directly into possible motor commands.[26]

More support for the vision-for-action hypothesis is the finding of the so-called 'canonical neurons'. In studies with monkeys, canonical neurons were the group of neurons discharging when the monkey observes three-dimensional visual stimuli which were congruent in size and shape with how the observed hand was shaped (Rizzolatti et al. 1988)**.** These neurons are considered to be involved in transformations of visual stimuli into motor command and thus are supposed to be necessary for object-directed

[26] This feature of visual processing in the dorsal pathway is most close to Gibson's (1986) idea of 'direct affordance pickup', however, as it is still regarded as representation of action-related features, thus crucially involving mental operations, it deviates from Gibson in at least this aspect (cf. Jacob & Jeannerod 2003).

actions (Jeannerod et al. 1995). This finding strengthens the claim that cognitive mechanisms can be determined even on the neuronal level responsible for action-related representation – by processing the visually perceived information regarding possible interactions.

Both the dual-systems theory and canonical neurons describe mechanisms for simple actions towards objects such as reaching, grasping or pointing. For all more complex actions, the direct relation of object feature and physical property no longer holds (such as hand span to handle size), which means that more complex actions can no longer be simply mapped onto simple features of objects and that memory, skill learning and experience determine the representation of higher level action possibilities.

8.2.3 Mirror Neurons

The role of mirror neurons (Rizzolatti et al. 2001) for human cognition is not entirely understood yet (see Rizzolatti & Sinigaglia 2010 for an overview). However, there is an ever growing body of evidence that clearly suggests that the mirror neuron system plays a crucial role in action cognition and action understanding. In contrast to the canonical neurons mentioned in the section before, mirror neurons would not discharge when the subject is presented with a three-dimensional object, but only when the subject either performs an action or when it observes a similar action performed by someone else.

The mirror neuron system also plays a role in the imitation behavior of neonates (Meltzoff & Moore 1983). Imitation implies processing visual input to subsequently generate a matching movement response. Being presented with different facial expressions, neonates showed imitation of these facial expressions. The conclusion from these findings is that there is a cognitive system, most likely innate or active from the very early developmental stages, that encodes visual input in terms of the subject's own motor activity. The mirror neuron system has been described as enabling this fascinating ability of neonates (Simpson et al. 2014).

This evidence suggests that one of the functions of the mirror neuron system might be the encoding of movements. Other studies, however, provide evidence that this is not the central and most interesting function of the mirror neuron system. The mirror neuron system is held by many to encode action rather than movement and thus also encodes intentions and action goals, which are not restricted to typical movements (Rizzolatti et al. 2000). As already mentioned, the mirror neuron system is supposed to not only encode the subject's own actions but also the observed actions of conspecifics (including the experimenter). For instance, it could be shown that the mirror neuron system encodes specifically goal-directed movements, instead of movements per se, e.g., in the case of an observed movement that is not clearly goal-oriented, such as hand movements imitating the actions towards an object that is absent. In addition, mirror neurons have been found discharging during goal-directed movements such as the flexing of a finger for grasping an object, but discharge weakly or not at all during the execution of similar movements that compose a different motor act such as scratching (Rizzolatti et al. 2000; Sinigaglia 2010). Furthermore, in a study with macaques, it was shown that the same group of the macaque's mirror neurons would discharge when both a grasping action with normal pliers (requiring opening the hand and then closing it to grasp an object) was performed and when reverse pliers were used (that required the unnatural movement of closing the hand first and then opening it to grasp the object). This suggests that it is not the movements that are encoded, but rather the action as such. In both cases, it was a grasping action which was encoded by the mirror neuron system, although in the second condition, the movements involved where rather untypical (Sinigaglia 2010; Rizzolatti et al. 2001; Rizzolatti & Sinigaglia 2010).

The mirror neuron system thus seems to be crucially involved in representing goal-directed actions performed by the subject or observed in others. What can be concluded for the present discussion of action-related representation is that there is good evidence for a neuronal system whose main function is to track goal-directed action. Mirror neurons would only discharge when one's own movements involve a goal orientation, and they would also discharge when the movements of others involve action goals. Representing the environment in terms of goal-directed action is thus even localizable on a neuronal level. It also helps explain why it is relatively easy for healthy subjects to immediately make sense of other subject's actions. Last but not least, with monkeys, primates and humans clearly being social animals, the discovery of a neuronal system that enables them to interpret goal-directed movement fits perfectly with the evolutionary significance of understanding other subject's doings, and for determining action-goals for oneself. From that perspective, the existence of a mirror neuron system strongly supports the claim that action-related representation is a very basic and important way of representing one's environment in terms of what can be done, for the subject and the subject's conspecifics.

8.3 Foundational Aspects of Action-Related Representation

Considering that action-related representation is supposed to be the grounds for further cognitive development and the development of representation with a more complex structure and conceptual representations, the question arises what the grounds are for action-related representation itself. Which processes or mechanisms have to be at work in the cognitive system of an organism that possibly could enable action-related representation?

A possible answer is provided by Bickhard (1999; 2009b), who claims that future-oriented anticipations of action-outcome are the basis for representation, but as has been shown previously, it still an open question if Bickhard's account of interactive representation can actually account for foundation of representation in general. Piaget (1954; 1977) suggests that the basis for the development of representational development is goal-oriented motor capabilities and the organization of their use and the received feedback. The idea presented by Piaget has been introduced in a similar fashion by O'Reagan & Noë (2001) under the term 'sensorimotor contingencies', originally meant to explain phenomenal conscious experience with motor skills and the feedback loops involved – which is largely irrelevant to this discussion. Nevertheless, their idea captures an essential aspect of movement and interaction that can be used to describe the basis of action-related representation.

Movements, despite their subjective aspect of being controlled by an individual, always have an objective aspect as they are closely linked to the environment. This close link always involves a feedback from the environment. Executing a movement involves both proprioception and exteroception, i.e., perception of (aspects of) the environment and of the subject's body (cf. Merleau-Ponty 1945/2012; Gibson 1986). Perception of an object causally depends on the movements a subject performs, and will change systematically as the subject interacts with it, (excluding that the object might change independently, as in melting away or dissolving or growing, etc.). This reciprocal connection of object perception and movement (perception) results in a *systematic covariance* between proprioception and exteroception: the perceptual apparatuses will receive the same kind of feedback from the environment given the same kind of movement. This sums up what O'Reagan and Noë (2001) refer to as 'sensorimotor contingencies'.

The information conveyed by these sensorimotor contingencies establishes the basis for the detection of simple action-related features of the world. For instance, a baby lying in a bed receives constant proprioceptive feedback from every movement. This feedback informs the baby about the solidity of the bed, or the mattress. If the baby now randomly, involuntarily or voluntarily, moves one of her legs and the leg would extend over the edge of the bedframe, the proprioceptive feedback the baby receives from this movement would be different from when it was all supported by the mattress. The baby would gather information about the consequences of that movement, such as absence of solid surface means being no longer supported, that there are edges on the bedframe and so on. The baby does not have any concepts for these experiences at first, so she will just recognize a change in feedback. This change in feedback will be systematic with some movements and so the baby learns that the bed has an edge beyond it no longer supports surface. With repeating random movements and processing the proprioceptive feedback, systematic covariance between movement and feedback is established from the earliest developmental stages on.

It is on this basis of the detection of sensorimotor contingencies that animals learn about the consequences of their movements and actions and thus establish action-related relations to objects in the world and to their environment in general. The first action-related representations of (objects in) the environment are formed simultaneously with sensorimotor contingencies detection on the basis of proprioceptive feedback processing. Another way to frame it is to say that the detections or registrations of sensorimotor contingencies are action-related representations, in the sense that self-generated movement is linked to feedback from the environment. The feedback from the environment establishes the object with its action related properties. What the subject learns at this stage is that if she had been confronted with an object with different action-related properties, she would have, ceteris paribus, perceived different action-related properties, and thus a different object. This is due to the fact that the sensory (proprioceptive) feedback would have been systematically different between the two instances, and thus the systematic contingencies between proprioception and exteroception would have also been different. Thus, proprioceptive information is a constitutive part of action-related representation development, such that a subject can only represent those action-related properties for which it has the according proprioceptive information. A subject can only perceive simple action-related properties (being a graspable thing, being solid, being within reach, being a thing that can be put in one's mouth, etc.) that are linked to movement that the subject is able to produce or has already produced.

Piaget's (1977) and Bickhard's (1999) accounts consider action to be the grounds of knowledge representation and cognitive abilities, with interactive representation being the foundation of representation in general (see ch. 7). What this implies can be made explicit by merging their ideas with the idea of sensorimotor contingencies as described above. To detect or register sensorimotor contingencies as described above, one does not have to possess and deploy concepts, nor does the detection and processing of sensorimotor contingencies necessarily have to be representational itself. This describes a development based on movement, which in most cases started as involuntary and uncoordinated, and subsequently coordinates itself while integrating sensory and proprioceptive feedback. Thus the first action-related representations consist of a tight coupling of self-generated movement with proprioceptive feedback, which allows for a progressive integration of more modal specific information. Defining action-related representations this way emphasizes the primacy of action (as self-generated, goal-directed movement) and proprioceptive feedback in contrast to visual information for the development of action-related representations. Thus, in this interactive picture of representational development, visual features of objects are merely features among other, action-relevant features that are represented in terms of movements. The available visual information will be integrated in the course of development in organisms, and some visual features might even have a special importance from the very beginning, as well as other sensory features such as smells that are important for the survival for many animals. Different ways of representing the world (e.g. representing the mother by detecting 'mother's scent'), having different representational contents, can coexist from the very beginning.

With this interactive approach in mind, representational content can be spelled out in terms of a combination of generated movements and the involved proprioceptive feedback relative to anticipations of goal states, which involve expectations of the state of the world at a certain future moment as well as the future state of the subject's body (cf. Allen & Bickhard 2013). By basic-level operations, in analogy to self-organized clustering in unsupervised machine learning (cf. Ultsch 1993), representational structures form on the basis of recognized structures of the environment, resulting from interaction. This way, simple classification of objects and other basic cognitive abilities can be accounted for with the interactionist framework, with the advantage that these basic mechanisms, which enable these processes, are by itself not representational but allow for representation generation. Of course, this is merely a sketch of the possible aspects of representation development and lot more of factors might contribute to forming complex cognitive abilities. However, these sketchy remarks illustrate the potential of explaining some aspects of cognitive development on the basis of action-related representation.

A further point concerns the scope of action-related representation: Action-related representation by no means has aspirations to be the exclusive foundation of *all* representation, such as Bickhard (1999) would claim. Thus, it is about elaborating the special characteristics of action-related representation, which is representing (action-relevant) features of (objects in) the environment in an action format – in terms of possible movements. In analogy, the special characteristics of visual representations would consist in representing visual features (of objects; the environment) in a visual format – a format that allows e.g. for visual discrimination of different objects by shape or color. Action-related representation enables goal-directed action in turn. Hence, the visually perceived (represented) redness of an apple might play no role for any interaction at all, while the representation of the apples size allows for detecting the grasp-ability of the apple and can thus be counted as an action-related representation. In other contexts, this distinction might become increasingly blurry, and a purely visual representation might turn into an action-related one. The general claim is that there is a whole variety of representations but they are not informationally encapsulated – information from the different modalities can be integrated in all sorts of representations, action-related representation being a paradigm example for the integration of proprioceptive and visual information in accordance with motor plans. Representing in an action-related format is already integrating information from different 'channels', while purely auditory representations are encoding only information from one 'channel'.

8.4 Summary

Action-related representations are representations of environmental features in an action format. Thus, they represent object features in terms of possible, goal-directed, object oriented movements. Representing possible movements of the representing subject, (basic) action-related representations are essentially egocentric. Basic action-related representations are the foundation for object representations by representing possible action-bundles that an object enables and the subject has learned. Action-related representation crucially involves information stored in the body schema, as physical aspects of the representing subject determine the range of possible interactions with the world. Their influence is mutual: experience and learning changes the set and quality of skills a subject possesses, which flows back into the body schema and enables representation of new action possibilities in turn. There is plenty of evidence for neuronal systems whose function it is to encode action-relevant information and directly represent it in an action format. Evidence for the mirror neuron system supports the claim that the distinction between mere movement and goal-directed action is already present at the neuronal level. Action-related representation can explain behavior and abilities of a variety of animals and human babies, hence action-related representation occurs in the absence of conceptual cognitive abilities, but may well be involved in developing conceptual representations and linguistic skills.

9 Development of Abstract Concepts

The general account of action-related representation developed in the previous chapter offers an adequate explanation for the underlying cognitive processes and states involved in subject-world interaction. In this chapter, its applicability for other cognitive domains will be demonstrated, resulting in a refined model of action-related representation that exemplifies different levels of complexity, gradually transitioning from the most basic action-related representations to more complex and sophisticated action-related representations. Having established this gradation model, I will show that action-related representations with differing levels of complexity exhibit different degrees of abstraction. Action-related representation thus is not limited to explaining behavior of animals, but can also explain the development of abstract cognitive abilities. As has already been claimed, action-related representation does not presuppose concepts and can thus be found across a vast variety of species and from early developmental stages on in humans. The capability for this kind of abstraction is therefore not limited by mastery of concepts and language, but rather by how developed the subject's skills are and what range of possible actions can be executed to which level of complexity. Abstraction mechanisms already occur on the level of simple interaction and develop with increasingly complex interactions – the capability for abstraction will then be used for other cognitive processes, such as the development of object concepts, which can be understood as abstraction from single encounters with the action-related properties of individual objects.

Interestingly enough, there is not too much work focusing on cognitive abstraction and the underlying mechanisms from a developmental point of view. Influential developmental work, such as Piaget's (1954; 1977) is a accounting for abstract cognition mainly by describing what happens at which stage of development. Thus Piaget describes when abstract cognitive abilities are likely to be developed and what other abilities this implies. Piaget is not explicit about the actual abstraction mechanism at work and on what basis these operations develop and what exact role they have for our understanding of cognitive abstraction.

Contemporary work on the development of abstract cognition, such as Dumontheil (2014) or Nee et al. (2013) focus on an entirely different aspect of abstract cognition, which is self-generation and stimulus independence. Thus, they take abstract cognition to be either thought about future or past events or goals, or to be about relations between representations rather than information processing regarding actual occurring stimuli. Dumontheil further mentions metacognition as an important aspect of cognitive abstraction. These aspects are all important in their own right and plausibly feature in higher-order cognition such as logical-mathematical reasoning or other highly creative thought processes. Nevertheless, they are not relevant to this discussion, as the central aim is to look for plausible abstraction processes on a non-conceptual level and Dumontheil focuses on development at later stages that already involve language and thus accounts for other kinds of abstract cognitive abilities (such as mathematical cognition)

The account focusing on based on action-related representation, which is presented in this chapters, deviates from other accounts of abstract cognition in that action-related representation describes abstraction mechanisms on a much earlier level of development, involving and presupposing no other sophisticated cognitive abilities. Furthermore, abstraction is described in relation to very concrete practical abilities, arguing that abstraction is grounded in the cognitive abilities that enable animals as well as humans to interact in goal-directed ways with their environment, thus providing an account of abstraction mechanisms that is adequate regarding evolutionary development and functions.

9.1 Basic Level Action-Related Representation

The analysis of abstraction mechanisms starts with the most basic kind of action-related representation. In chapter 6.2, while analyzing Jacob & Jeannerod's (2003) account of lower and higher level pragmatic processing, it became clear that there must be a distinction between different levels of complexity in representing action possibilities. Hence, the most basic level of action-related representation, following their analysis and the general account of action-related representation developed in chapter 8, is comprised of a direct 'mapping' or 'matching' of salient object features and elementary properties of the subject's body. Examples are:

> graspable things like bottles, handles or clutches, where the object's diameter is mapped onto the subject's hand span;
> step-on-able things like stairs or chairs or steps, where the object's height is mapped onto the subject's leg length;
> reachable things like objects in the subject's immediate surroundings, where the object's distance is mapped onto the subject's arm length.

These representations relate the physical properties of an object to bodily properties and thus enable interaction: what is represented as being less than an arm's length away is potentially reachable. The claim is now that this way of representing objects as 'within reach' is a primitive way of representing distances, in that the understanding of distances in terms of reaching space is purely egocentric. This primitive representation of distance could well be the first way of representing distance at all; Moreover, this way of representing distances is clearly action-related. The primitive representation of an egocentric distance is also an action-related representation of reachable objects in one's personal reaching space. The representation is not yet one of complex structure: it neither has to explicitly represent the subject and the object as such, nor the distance in an explicit way (such as 'an arm's length away'). It is enough that an object is represented in an interactive way, as an object that can be taken possession of by reaching out and grasping it. The object, and thus the distance, is represented by a simple action. This implicit way of construing the representation of features of one's environment is analogous to Strawson's (1959) notion of 'feature-placing'. Feature-placing sentences do not pick out an object by describing the object and its properties. Rather, the feature-placing sentence asserts that a certain feature is present, as made clear by the example: 'it's raining'. Here, it is not stated that 'there is a time and a place' to which the property of 'rainy-ness' is attributed by a propositional sentence, but rather that the feature of 'rain is happening' is ascribed to a situation which is neither further described nor characterized – thus only, if at all, implicitly represented (see also Chemero 2003). Feature-placing refers to situations, not to objects – this is the way one should also think of the basic action-related representations such as 'within-reach'. This is what Campbell (1993; 1994) had in mind with his notion of causal indexicals, which captures the idea of implicitly representing environmental features that have immediate implications for action, by representing them as causally indexical. What is not part of the causal indexical is an explicit object-property representation, but rather a situational feature: '(is a gap that) is jumpable (for me)' where the parts inside the parentheses are implicitly represented – the other part is the explicit action-related content of the representation. The object ('the gap') does not have to be explicitly represented and neither does the egocentric aspect – it is clear that only the subject who is forming, entertaining and deploying the causal indexical representation is meant. It would be evolutionary disadvantageous for an animal if it were to explicitly attend to the subject (themselves) of an interaction possibility every time an interaction possibility is discovered. This would likely mean having to conduct an extra cognitive step at the beginning of any interaction. On top of this, it would also presume having the concept of a self that can be ascribed to. In this sense, causal indexicals are basic action-related representations, and are similar to Strawson's (1959) idea of feature-placing.

9.2 Intermediate and Higher Level Action-Related Representations

The next step in the development of more complex action-related representations consists of two aspects: First, overcoming the purely egocentric perspective, manifested in the implicitness of representational content, and second, explicitly representing objects and agents. This leads to a more general representation of action and is the basis for developing object-concepts – at least of those objects that allow for physical manipulation by agents. The development of object-concepts involves the categorization and clustering of action-related features of objects, derived from singular encounters. Repeated interaction with such objects forms action-related object-concepts. The second aspect consists of learning that agents are the subjects of actions – first, the subject has to develop an explicit understanding of herself as an agent, then this knowledge can be transferred to other agents. By applying this agent-related knowledge, the subject will be able to explicitly relate action-relevant properties of objects to other agents and even classes of agents. At the core of ability is then formation of action-related representations which are able to represent subject independent action possibilities. This is the highest level of complexity in the general account of action-related representation and at the same time exhibits the most abstraction: abstraction from the original agent, situation and individual encounters with objects. Describing cognitive abstraction mechanisms on the level of action-related representation provides a new approach to the problem of grounding abstraction. Whereas most other approaches mainly focus on cognitive abstraction processes on the basis of perceptual representations (cf. Barsalou 1999; Prinz 2005), the proposal in this chapter is to consider action as the starting point for cognitive abstraction mechanisms. As action is defined in terms of goal-directed movement, cognitive abstraction can be understood as grounded in movement of a subject's body – providing a new angle for the embodied perspective on cognition.

9.2.1 Implicit and Explicit Representation of Objects and Agents

The most important aspect of the cognitive abstraction mechanisms being grounded in action-related representation is the transition from egocentric representation of action possibilities to an allocentric representation. This corresponds to the distinction made by causal indexical representation in contrast to representing action possibilities for other agents. On the one end of the abstraction process, subjects are only able to implicitly represent an action possibility, as described by examples such as 'within reach (for me)', 'is a weight I can easily lift' etc. On the other end of the spectrum, we find representations of possible actions for other agents: 'Is only within reach for a tall person', 'is a lid (e.g., of a cucumber jar) that can only be opened by adults' or 'can only be walked through by subjects smaller than 1,50m'. At the core of this transition lies the distinction between implicit and explicit representational content. Only by gradually developing an ability to explicitly represent certain action-relevant features, the further transition from egocentric to allocentric can be achieved. In a way, these two aspects are mutually dependent, but will be treated separately here for the sake of clarity.

In basic action-related representations, subjects do not need to explicitly represent all information relevant for the interactions success.

Dienes & Perner (1999, 736) state: "a fact is explicitly represented if there is an expression (mental or otherwise), whose meaning is just that fact". In other words, explicitly represented information is an analyzable part of the representation.[27] In a proposition such as 'this is Johns car', 'John' is an explicit part of the propositional content. Furthermore, John also happens to be the neighbor of the person expressing the proposition, thus the person is also referring to her neighbor, alas only implicitly, as 'neighbor' is only an implicit part of the proposition's content. Another example is 'John is a bachelor', where bachelor is an explicit part of the content, that can only be fully understood by the implicit content 'bachelor = unmarried + male'. Thus, information is implicitly represented, if it is relevant for the meaning other explicitly represented parts have for the

[27] See Dienes & Perner (1999) for a discussion of explicit and implicit knowledge.

subject, but is not itself a part of the explicitly represented information. (cf. Dienes & Perner 1999).[28]

In the same way, an action-related representation can have implicitly and explicitly represented parts/content. Basic action-related representations, as captured by causal indexicals (Campbell 1993; 1994; see ch. 4) represent action-possibilities in a given context, such as 'this is within reach (for me)'. This representation refers to an object or an environmental situation as well as referring to an agent – though only implicitly. The explicit content consists of a representation of an action possibility, but at the same time it is quite obvious that the representation must also contain at least an implicit reference to the agent and the features of an object, in order to provide an explanation of behavior at all. Without the implicit reference to the agent and an object, it would never be explainable why a specific subject entertaining this representation would act, representing only an abstract, detached action, such as 'reaching'. In representing action-possibilities, subjects always represent the possibilities for themselves as possible agent, which is due to the egocentric format of the representation. As discussed in chapters 4 and 8, egocentric representations are already representations in an action format, representing the possible action space of subjects. Moreover, the subject needs to represent an action goal, which in many cases involves an object towards which the subject will direct her action. This action goal needs only to be implicitly represented as well and does not have to be an explicit representation of an object as the object, such as being directed towards a cup and representing the cup as a cup. Although this is what normally happens, the implicit representation of an object as the referent of 'within reach' is enough to explain and guide a subject's action towards this object.

Implicit representation, as construed above, does not presuppose conceptual knowledge, or does so only to a very minimal extend. Representing an object in a basic action-related way refers to the object only by referring to the aspect relevant for executing the action. In the case of a gap, which is represented as jump-over-able for the subject, all that is represented is a possibility to successfully jump over a certain distance. The

---

28 Perry's (1978) notion of 'unarticulated constituents' is of the same spirit. (See also ch. 4)

absolute distance is of no importance at all while all sorts of animals have very good means of determining distances for actions without having any concepts of distance that can be ascribed to an object, i.e., the gap in this case. The representation also does not need to be permanent – it could be formed in a given situation and then be discarded as soon as the context changes. For actions to be successful, all that is needed is that the respective action-goal be adequately represented in the right moment – no stored representations are needed to do this. For a squirrel to adequately determine the jump-over-ability of a gap, recognition of the gap by reactivating former experience with gaps is not necessary. In dynamic situations, the squirrel will not rely on memory, but on situational action-related representation.[29] What the squirrel actually does is picking out an action possibility – as in demonstratively referring to an unknown object by pointing at it and saying 'this'. The action-related representation thus construed picks out a situation by enabling a certain kind of action. A feature of an object is thus represented in terms of action – a possibility for a jumping action represents the gap.

9.2.2 Transition from Egocentric to Allocentric

Explicit representation is already a form of abstraction. The abstraction involved in explicitly representing an object as object is characterized by generalization and classification. This step crucially involves memory. Repeated interaction and different interaction types with an object lead to representation of an object with various aspects. The idea is similar to Allen & Bickhard's (2013) interactive approach to object representation (see ch. 7.2). Allen and Bickhard's action-based model of object representation claims that representing an object is the same as knowing how to competently interact with the object (cf. Allen & Bickhard 2013, 128). Furthermore, every object offers multiple interaction possibilities that are

29 Squirrels do seem to rely on memory when relocating buried food much more than initially assumed, as has been shown by Jacobs & Liman (1991). These findings support the claim that squirrels use spatial memory and not solely olfaction for finding buried food. Squirrels thus may use in fact cognitive maps and rely much less on dynamic aspects - another reason to generally embrace representationalism for all sorts of animals.

grouped and will form a web of interaction possibilities which will constitute the object representation. Every object in this sense is represented as an invariant organization of interaction possibilities. In addition, perceptual information becomes integrated with the functional information about the interaction possibilities, from which a fully-fledged object representation emerges (cf. Barsalou (1999) and his notion of a 'simulator'). Important at this stage is that most traditional accounts focus on the perceptually available information and neglect the information from interaction. The aspect of interaction that is of importance is that it consists of output and input conditions at the same time, while purely perceptual representations are understood by most in a more passive way, consisting of sensory input and cognitive processing. The interactive approach crucially involves the movement output as well as proprioceptive feedback and anticipation of future states.

Generalization and classification is the process of learning from repeated interaction. As the behavioral repertoire of human babies constantly develops and allows for increasingly complex manipulations with objects, the feedback information becomes more complex too. Developing more complex object representations are thus bound to developing more complex behavioral skills. At the same time, cognitive development proceeds, and abilities such as recognition and memorizing more complex events and situations improve and a sense for object permanence develops. These basic cognitive abilities all contribute to the development of explicit representation of objects and object features. Piaget (1952)[30] found that a grasp of simple objects develops at around 8 months. Children at this age would start to search for an object if it moves behind an occluder, implying that they have awareness or the expectation that the object is still existent, while children younger than 8 month would simply show indifference regarding the object once it disappears behind an occluder. The capacity for full object permanence develops at about 18 month (see

30 Whereas Piaget's (1952) focus was on the development of general cognitive abilities, the development of action-related representation is about specific representations and how they emerge. The reference to Piaget thus should be understood as presenting an analogues explanation for cognitive development and applying it to the development of more abstract representations on the basis of interaction.

Rakoczy 2010 for an overview). Object permanence is a crucial step in developing explicit object representation, as it implies that objects are recognized on the basis of their features, and thus is different from the feature-placing that happens in the most basic level of action-related representation. Feature-placing (cf. Strawson 1959) is held not to involve any kind of object persistence and is thus the most basic mode of object perception, as it consists merely in representing and reacting to present features, which can be done without any reference to a definite object. For explicitly representing an object, the infant thus at least must be able to spatio-temporally track objects and understand objects as persistent things that are independent from the infant's existence. The crucial cognitive development appears to happen already on a pre-linguistic level, but even more so at around 10-12 month, where language development also starts (cf. Xu & Carey 1996).

The transition from implicit to explicit representation also takes place at the level of representing the agent in the interaction possibility. To refer explicitly to herself, a subject must have a concept of herself as a self. This involves being able to recognize oneself as oneself as well as being able to recognize others as agents, to finally discovering that others have mental states such as beliefs and desires and act intentionally. The different stages in early childhood development that allow for these abilities are subject to research and a complete overview cannot be given here. For non-linguistic infants, the standard test for rudimentary awareness of oneself is the so-called 'mirror-rouge' task (Gallup 1979), where a red mark is attached to the infants' foreheads without their knowledge, then being confronted afterwards with their mirror image. The infants reaching for the mark on their own face are taken to have a sense of 'selfhood', that they are an object in space. Only children from 18 month onwards and great apes have demonstrated mastery of the task, while younger children and many other species regularly fail the test (cf. Tomasello & Call 1997).

Evidence that others are represented as beings with mental states is provided by studies such as the famous false-belief task (Wimmer & Perner 1983). In these experiments, it could be demonstrated that only children from 3-5 years of age are able to take the perspective of others and attribute mental states, at least in a rudimentary form. The infants of the experiment had to decide where another agent would look for, e.g., a toy that had been relocated without the knowledge of this other agent, though the test infants could see where it was hidden. Below the critical age of about 4, most infants consistently fail to master this perspective taking and assume that the other child will look where the toy has been relocated, relying only on their own knowledge but not considering that the other agent lacks this kind of knowledge, hence has other beliefs about the location of the object.

From this it can be concluded that developing a notion of 'self' is prior to thinking of others also as 'selves', and thus contributes to a general concept of agency and agents. The ability to represent other subjects as possible agents of an action comes along with the transition from egocentric to allocentric action representation. This idea is captured by Synofzik et al.'s (2008a; 2008b) 'two-step account of agency', in which a distinction is made between 'feeling of agency' and a 'judgment of agency'. In further analyzing the notion of 'self-consciousness', they distinguish the 'sense of agency' and 'sense of ownership' as two core features of the notion of a 'self'. Thus, the 'self' is a complex representational structure that develops gradually, starting from non-conceptual implicit self-representation to gradually transitioning to a conceptual level and resulting in meta-representational abilities. Most interesting about their proposal is that they identify the basic level with a "sensory registration of action–effect-couplings" (Synofzik et.al. 2008, 413), thus considering basic sensorimotor processes as starting point for the development of self-consciousness. The feeling of agency is a product of perception based representations of agency, involving mainly proprioceptive and sensory feedback and is an implicit self-representation. The feeling of agency tells the subject only if an action was caused by the subject, by processes associating the motor commands with proprioceptive and sensory feedback. The perceptual representation of agency is non-compositional and does not have a property object structure (cf. Newen and Bartels 2007; Vosgerau 2009). Thus, this primitive self-representation is based on early and basic interactions and enable the further development of the ability to form 'judgments of agency', which is based on conceptual representations and propositional thought, allowing for referring to oneself as the agent of one's actions. (cf.

Synofzik et al. 2008a). The feeling of agency is still present in judgments of agency, as a representational core the judgment builds upon. To attribute actions to other agents requires the ability to form judgments of agency, the primitive core feeling of agency does not allow for attributing this feeling to other agents – a subject cannot have a feeling of agency for other subject's actions. Thus, the notion of 'self' develops on the ground of interactions that imply perceptual feedback and give rise to a sense of agency that is crucial for a 'self'-concept, which in turn is a prerequisite to think of other subjects' actions and is thus an important step in the development of cognitive abstraction.

The first action-related representations are by definition egocentric, as they are only used to guide the subject's own actions. Later stages of development enable children to acquire a general sense of agency and of others as agents. Representing others as agents is constituted by being an agent oneself and being able to have an explicit representation of one's own agency. A basic action-related representation, such as expressed by a causal indexical term ('this is within reach'), does not specify the agent explicitly as no agent at all is needed to be referred to in order to successfully reach for it or to move ahead as long as an object is within reach. When the child develops a sense of agency in general, she will also be able to think about other subjects as agents. Thinking of other subjects as agents involves action-related representations that relate bodily properties or skills to properties of objects. A simple example would be a child that wants to reach for the cookie jar on the kitchen counter but is too small to actually succeed. The cookie jar is not 'within reach' for the child. From a certain age on, the child will signalize (by pulling the adults arm, by pointing at the jar, etc.) to grown-ups nearby that she wants them to get the cookie jar for her. This behavior can be explained with an action-related representation that represents an object of desire and an agent, which is not the child and whose abilities are put in relation to contextual features. The child's representation has the content '(the cookie jar) is reachable for mommy'. Action-related thinking as described in this scenario is clearly more abstract than a child discovering reach-ability for herself only, by looking at an object in her vicinity and grabbing it. The next step in the development of more abstract action-related representations is to represent possible actions for whole classes of individuals. Typically, children will learn that some things can only be done by adults, e.g., their parents. In that sense, they will develop action-related representations that have 'adults' as a part of the content, such as 'the apple can only be cut by an adult (handling the knife)'. The important aspect here is that the ability to think about other people's action capabilities, as well as what objects afford for other agents, has to be learned – it is not in the cognitive repertoire from the beginning and infants and young children most certainly have no, or only limited means of representing these action possibilities.

Among the possible grounds for developing the ability of thinking about action possibilities for other agents is most likely the improvement of one's own behavioral repertoire and set of skills. The more a child learns about tool use and manipulation, about using one's own body in more sophisticates ways ('doing cartwheels', learning how to tie shoelaces', 'learning how to swim or ride a bike') changes the individual's body schema significantly and opens entirely new kinds of feedback responses that provide new information, which can again be integrated in the existing information about one's body and the world. By successfully managing to execute newly acquired skills, the child will acquire a new understanding of movement. When just watching other subjects perform movements the child has never encountered nor performed herself, the understanding of what exactly is happening will be a limited one. Only by reenacting and developing similar skills, will the child be able to get a better grasp on observed behaviors of others. In this sense, it is only natural to claim that the first action-related representations are of limited complexity due to the set of performable skills of infants, and become increasingly more complex over time. The ability to represent more complex action possibilities is the condition for representing action possibilities for other agents explicitly and no longer only for the child herself.

To avoid misunderstanding, this is not meant to claim that all subjects can only understand actions of others that they are actually able to perform themselves, which would be too strong a claim and is not backed by empirical observation. One can of course reason about football game strategies on the highest level of abstraction possible, without ever reaching a professional level of playing competence. The claim is rather that developing skills is of importance in the early years of life and enables transfer to other situations and agents. Plausibly, having a wealth of personal experience in many respects will help in developing richer concepts of the skills involved, but this is not to say that the actual reenacting is a necessary or even sufficient condition to understanding actions.

The more a cognitive system develops, the more conceptual representations and abstract inferential skills the system has, equating to more ways of acquiring knowledge – by way of cross-experiential transfer and inference on the basis of past events. Perhaps the most promising way of accounting for cognitive processes grounded in motor skills is to formulate a moderate thesis that allows for constitutional grounding as well as decoupling from the original experience. Weber & Vosgerau (2012) do provide a good argument that the thesis of 'grounded cognition' can be best understood if it is constitutional to some degree, but not necessary for all cognitive processes. Some cognitive abilities develop on the grounds of motor abilities, where the latter can become lost without affecting the first – at least not significantly. Other cognitive skills can become impaired when motor skills are impaired too, showing that there are constitutional relations between cognition and action. Thus, some motor abilities are constitutional for some cognitive abilities, but higher level, abstract cognition can become independent from motor abilities in the course of both ontogenetic and phylogenetic development (cf. Weber & Vosgerau 2012, 62). They present a number of studies that provide evidence that impaired motor control leads to impairment (but no complete loss) of cognitive abilities. Patients suffering from amyotrophic lateral sclerosis have shown deficits in word-description matching associated with judgments about actions (cf. Grossmann et al. 2008). This demonstrated a connection between motor deficits and knowledge of action features, suggesting that thinking about action is indeed connected to being able to act. Furthermore, patients suffering from Parkinson's disease showed drastic impairments in action naming tasks, while performance was comparatively better in object naming tasks (cf. Rodríguez-Ferreiro et al. 2009). A study by Bosbach et al. (2005) involved two deafferentiated patients, who have no proprioceptive feedback anymore and rely on visual feedback for motor control. The patients (and a control group) had to evaluate another person's anticipation of weight, observing them lifting boxes. The patients had significant problems deciding whether the observed action of other agents was correct or if the action preparation did not match the actual weight of the box and therefore too little or too much force was applied when attempting to lift the box. The control subjects had no problems telling that the agent thought the box was heavier than expected and therefore prepared for a stronger lifting-movement, which results in the box going up too fast, together with a reaction of surprise. These studies provide evidence for the claim that impairments in motor abilities can cause deficits in formerly developed cognitive abilities involving action-representation, although no complete loss of action cognition could be shown (for details, see Weber & Vosgerau 2011; 2012).

For the present discussion, these findings support the claim that some aspects of the development of cognitive abilities are grounded in developing motor abilities. This justifies the claim that on the basis of developing skills for action, children thus acquire other cognitive abilities, such as representing more general action possibilities for themselves and other agents. Abstraction in this sense is achieved by abstracting from oneself as the only possible agent and being able to represent all sorts of possible actions regarding other agents. This also implies abstraction from one's own body schema and other physical properties and relating the features of objects to possible agents in terms of possible movements – the subject mastering this level of abstraction represents possible movements for all sorts of individuals in relation to objects and their action-relevant properties. The abstraction consists in explicit representation of agent and object, extending the frame of reference from purely egocentric to a detached allocentric frame and finally in perspective-taking – thinking of action possibilities for other subjects and even of classes of subjects.

9.3 Developing Object Concepts on the Basis of Action-Related Representation

As has already been indicated in the previous paragraphs, explicitly representing the object involved in possible interactions is crucial for a more sophisticated mode of action-related representation, together with a higher degree of abstraction. This abstraction process is involved in the development of other cognitive abilities, such as conceptually representing objects. Action-related representation, from the most basic to more complex levels, provide the foundations for the development of object concepts – at least of those objects that have a meaning for interaction (which is true for the majority of objects)[31]. The cognitive mechanisms underlying conceptual object representation involve classification and generalization. Both cognitive operations are abstractions from the specific to the general - in the action-related framework, this implies abstraction from specific individual interactive encounters with objects to a representation of an objects functions in general, or even an object without explicitly representing the functions anymore. The individual encounters with objects can simply be based on basic action-related representations, but more experience and the learning of invariant features leads to a more complex, generalized representation of objects, which is finally the basis for conceptual object representations. The transition is thus from singular encounters with, e.g., round, inflated, kick-able objects to the general concept of ‘ball’. The basis for this object-concept development does not, as has been widely held, occur on a purely perceptual basis (cf. Barsalou 1999), but crucially involves interaction and thus action-related representation.

[31] This claim is clearly true for artifacts, but not without restriction for natural entities. The action-related properties of trees will probably be less salient and therefore less frequently produced by many subjects that those of cups and hammers, which were designed with a default function. Still, trees are climbable for some subjects who therefore are quite likely to have a more action-based concept of trees as compared to non-climbers. This again highlights the role of previous experience and skills in representing the action possibilities of one’s environment, contra Gibson (1986) (see also Borghi 2004).

It can be demonstrated that subjects represent objects in terms of their salient functional properties or what interaction they allow for, rather that other feature that are of no obvious action-relevance (e.g. the color of an object). Borghi (2004) provides evidence that objects are represented as patterns of actions. In one study, participants had to name component part of complex objects (e.g. bicycle, piano, and mixer) in three conditions: seeing, constructing or interacting. In the interaction condition, the focus was on specific parts, which are highly relevant for using the object. In general, action-related parts were produced more frequently and earlier than non-action-relevant parts. Borghi (2004) concludes from these and other findings that objects are represented componentially rather than holistically and the components with functional relevance are represented more frequently in interactive situations. Moreover, in varying situations, different action-relevant properties of the objects were produced by the participants, suggesting that affordance representation is task-related. Representing objects thus means representing action possibilities, depending on the situation, leading to the general conclusion that object concepts are action-based (cf. Borghi 2004, 23). Furthermore, in a study by Iachini et al. (2008), it could be shown that object features, which are action-relevant, play a significant role in categorization. In all experiments, the property most relevant for interaction was considered to be the most important property for categorization, particularly in sorting tasks. Magnié et al. (1999) were able to show that action is a powerful cue in recalling information about objects. A patient suffering from severe object agnosia could recognize only objects with which he could recall associated actions – tools, kitchen utensils, clothes etc., but no musical instruments (the patient never played any), and also had problems recognizing animals. The patient could recognize actions and was able to produce gestures appropriate to using certain objects. These findings suggest that sensorimotor experience is of major importance for the representation of some object categories, namely tools and other artifacts with a clear functional description. This sits well with the patient's troubles recognizing musical instruments, with which he had no significant sensorimotor experience.

All the evidence presented supports the claim that action is of great importance for the development of object concepts and actual object representation in cognition. The possible actions a subject associates with an objects is or becomes part of the representation, although, as has been mentioned in the previous section, the link to action can become lost. In this sense, the focus of action-related representation is on concept acquisition and the development of conceptual representation on the basis of action-related representation. Central to the claim that action-related representation can account for the development of conceptual representation is, once again, the idea of a gradual transition and hence an increase in complexity. On the one end, basic action-related representations, such as causal indexicals ('within reach'; 'too hot') can be described as non-conceptual representations of possible actions. Higher level action related representations, such as generalized interaction properties ('is for cutting for adults'; 'this type of box is lift-able for strong persons only'; 'violins are playable only for skilled professionals') can be interpreted as conceptual representations of possible actions.

Before analyzing the gradual transition from non-conceptual representations to conceptual ones, it is helpful to elaborate on the defining criteria for conceptual representations. For this purpose, Newen & Bartels (2007) provide a very useful list of criteria a cognitive system needs to fulfill in order to be justifiably attributed the possession and entertaining of conceptual representations. Newen & Bartels (2007) are mainly concerned with finding cognitively adequate criteria, which justify attributing conceptual representations on a behavioral basis, therefore not presupposing or implying any linguistic capabilities. Moreover, they seek to add cognitive significance to the well-known criteria of productivity, systematicity and compositionality (cf. Fodor 1990) as well as taking Evan's (1982) 'Generality Constraint' into consideration. The Generality Constraint requires of systems entertaining conceptual thinking the following characterization:

> if a subject can be credited with the thought that *a* is *F*, then he must have the conceptual resource for entertaining the thought that *a* is *G*, for every property of being *G* of which he has a conception. (Evans 1982, 104)

In Newen and Bartels' terms, this is spelled out by stating that a cognitive system possesses concepts only if it is able to discriminate different properties in an individual object and ascribe an individual property to different objects (cf. Newen & Bartels 2007). Thus, their account is an important contribution to the debate about necessary and sufficient criteria for concept ascription to animals or cognitive systems in general. The first criterion they define is that the representation has an *object-property structure.* This entails that the representation of an object attributes a property to this very object in such a way that both the object and the property are represented independently, i.e., either one object can be represented with different properties (e.g., 'red/green/blue → square') or the representations ascribes the same property to different objects (e.g., 'red → square/triangle/circle'). This could be rephrased that the representation has to have a structure that is in principle analyzable as components that can be represented independently and in different contexts. The opposite would be the representation of an object that happens to be a red disc, simply as 'this object' in a demonstrative way. This representation would still allow for discriminatory behavior, as in distinguishing 'this object' from 'that object', very similar to how simple pattern recognition systems or color or movement detectors work. A camera that can detect instances of 'red' might be said to represent something as 'this', but this representation is far from being conceptual – for the reason that no object-property structure that can be analyzed is underlying this representation.

The next aspect of conceptual representations, according to Newen & Bartels (2007), is *relative stimulus independence.* This means that one stimulus that caused, or triggered, a representational state, can cause another at a different time, e.g., a red square might activate a 'red'-representation at one instance and a 'square'-representation at another. Moreover, the conceptual representation can be activated in the absence of the characteristic stimulus, that originally lead to forming the representation, such as a 'red'-representation could be formed due to an encounter with a red object and be activated again by an acoustic signal. Newen and Bartels third criterion for a conceptual representation is its role in a *minimal semantic net.* This means that the contents of the system's representations are systematically related to each other, at least partly, and are thus allowing for a basic classification according to feature-dimensions: instances of 'redness' are represented as instances of color as opposed to shape or material. The minimal semantic net enables representing something *as* something, in a primitive way. These aspects taken together are the necessary and sufficient[32] conditions for conceptual representations. This set of criteria can be applied to the distinction made in this chapter between basic action-related representations and higher-level action-related representations, which reflects the transition from non-conceptual to conceptual representations.

9.3.1 Non-Conceptual Action-Related Representations

Following from Newen & Bartels' (2007) framework, basic action-related representations can be interpreted as are non-conceptual representations. This can be demonstrated with the example of causal indexicals, such as 'within reach' or 'too hot'. While a core feature of conceptual representation is that it represents the object independently from its properties – hence conceptual representations attribute properties to objects – a representation such as 'this is within reach' is an action-related representation that can be used to explain and guide the subject's behavior, but lacks the required internal structure of property-object attribution to be a genuinely conceptual representation (see also ch. 4)

What is attributed here is a feature in a context: the agent represents reach-ability in her environment and can do so without even representing the object within reach *as* the object it actually is, e.g., a bottle. Moreover, it is hard to identify the object of the property ascription, as the goal-object in the causal indexical is represented in a demonstrative way, as 'this'. Even if one would accept that a property is attributed to an underspecified entity, i.e., object-property attribution in a minimal sense, a causal indexical would fail to meet Newen and Bartels' (2007) other criteria. 'Within reach' is highly contextual, thus it is not sensible to assume that the representation would be formed or deployed outside of a situation where a

32 Though the criteria are sufficient conditions for conceptual representations only if they are all realized – taken on their own, they can establish only necessary criteria.

subject is actually confronted with something that is within reach. In fact, context sensitivity is a defining criterion for casual indexicals, hence the term 'indexical', which refers to varying reference in differing contexts. Furthermore, it is unlikely that the requirement of a minimal semantic net is already met on the level of basic-action related representation.

Basic action-related representations are supposed to explain simple interactions with one's environment in cases where the subject is confronted with situations that have a certain degree of congruence with the subject's bodily properties: situations where a subject is cycling and approaching a tree with a low hanging branch and the subject has to decide immediately whether to duck her head to pass under the branch. A bottle on one's desk that is within reach; a hole in a box where a hand might fit through – all these examples for basic action-related representations lack the kind of systematicity and stimulus independence that is required for having a minimal semantic net in order for one's representations to be in a systematic relation due to their representational content. 'Within reach' can be action-guiding but is not systematic and woven into a net of related contents. Moreover, causal indexicals are not suitable for classification (along property dimensions). A class of 'things, within reach' would not really comprise a helpful class of neither action possibilities nor objects, as it would not be applicable outside specific contexts anyway.

Altogether, basic action-related representations fail the requirements for conceptual representations due to the fact that basic action-related representations are mainly representations of possible movements without having a full-fledged object representation as part of their content. What is underlying a 'within reach' representation is a grasping movement as in 'this is only a grasping movement's distance away', and this is not a property proper that is attributed to an object. Rather, it is a way for the subject to spatially represent it's immediate environment that is constituted by the actions the special framework allows for. Hence, it would make no sense to analyze 'within reach' as an attributed property of 'this bottle'. For that reason, examples of causal indexical action-related representation can exist in relative isolation and still maintain their cognitive function as well as their desired explanatory role.

9.3.2 Conceptual Action-Related Representations

In cognitive development, basic action-related representations can be considered only as the starting point for representing action possibilities. The behavior of many animals, including humans, shows a degree of complexity requiring more complex action-related representations to explain the behavior. Following the gradual transition model, action-related representation of a certain degree of sophistication can be described as conceptual. These include complex actions, such as actions involving tool use, in which properties are attributed to objects. The subject considers the chair step-on-able, and of adequate height, so she can reach for the upper shelf. She also evaluates the lighter as a substitute for a bottle opener, being a lever that can decap the bottle – which she might have observed others doing. These examples illustrate the various and manifold ascriptions that take place in complex actions. To meet an intended action goal, a variety of means are defined as useful towards obtaining the goal. These objects are all ascribed action-relevant properties in order to function in a larger context and these action-relevant properties can also be ascribed to other objects. Many objects afford substituting a bottle opener, so the same action possibility is attributable to different contexts. Also, the attributions on this level are related to each other in terms of their basic properties that allow for certain actions. A lighter can be a bottle opener, but so can rulers or screwdrivers, because they share some properties, and once the skill of opening a bottle is learned with one of these objects, the skill is transferable to different objects that share their properties. Furthermore, the criterion of relative stimulus independence is given: action-related representation can represent novel possible actions and can ascribe new action-possibilities to objects, even in the absence of the objects. One can think about new ways of combining things for a purpose that has never been thought of before in this very way, which is the basis for practical problem solving and technical inventions.

In generalized action-related representations, it is even less complicated to show that the conditions for conceptual representation are met. Thinking about possible actions for other agents requires the ability to ascribe various properties to the agent and the objects involved, so that the representation is about physical properties and skills of other agents in relation to properties of objects. A chair that is for sitting only for adults, but for climbing for toddlers, or making the house child proof by thinking about the possible dangers that might occur due to possible actions of the toddler: where she could get stuck, fall down, bang her head, what dangerous items are within reach etc. These are examples of thinking in terms of possible actions for others by relating their physical properties to their environment, where the action-related representations are on such a high level of abstraction that they clearly fulfill the requirements for conceptual representations.

Action-related representations can be the basis for conceptual representations, while not presupposing conceptual representations. The previous examples of generalized action-related representations clearly involve concepts already – action possibilities have to become conceptualized as well as a solid basis of stable object concepts is required for thinking in terms of actions for classes of individuals. But this does not entail that possessing concepts is the general condition for all instances of action-related cognition. What is needed to reach the level of abstraction that is implied in representing actions for others are the cognitive abilities of object permanence and, even more important, a concept of self and the ability of perspective taking. Developmental literature tells us when these abilities are usually developed (see the ch. 9.2.2), and there is some evidence concerning the role of action in these developments in humans and animals. Piaget (1952) showed that the development of object permanence happens along with the onset of instrumental action, which is behavior that is structured by goals and means. Infants around 8 month of age show not only object permanence, but also clear goal-oriented behavior, such as removing an object in order to reach for another object and they also start using tools for an end with increasing efficiency (cf. Willats 1985). These findings show that goal-oriented action and object permanence are related abilities, and while it seems perfectly reasonable to claim that goal-oriented action presupposes object permanence, the opposite might also be true: that object permanence is learned by interaction with the world and learning the sensorimotor contingencies that interactions with different objects provide. At this stage, it is almost impossible to treat these abilities as separate and thus a mutual dependence and development is the most likely explanation. What these findings also show is that goal-directed action is established long before the ability to discriminate and recognize oneself from other objects develops, which happens around 18 month in infants (see ch. 9.2.2 on the 'mirror rouge' task). From a temporal perspective, goal-directed action is established earlier than the development of a 'self-concept' and also earlier than the ability for perspective taking, as demonstrated by the false-belief test, which is mastered at around 3-4 years (cf. Wimmer & Perner 1982). Anderson et al. (2013) provide evidence for the changes in perception-action coupling, spatial cognition, memory and social development that are all interpreted as consequences of the acquisition of independent locomotion. They can show that infants who have a delayed development of independent locomotion due to neurological or orthopedic causes also exhibit limited spatial-cognitive abilities. Spatial-cognitive skills in turn are needed for successful object-oriented actions which points to a central role for the development of motor skills from which other, cognitive and behavioral skills emerge. Sommerville & Woodward (2010) provide evidence that understanding the actions of other agents as goal-directed is a function of the infant's own action production. Piaget (1954) provided some interesting insights on the connection of locomotor abilities and cognitive development, too. According to his reasoning, the change from an egocentric perspective to an allocentric occurs together with developing locomotion. This is necessary because as long as the infant is stationary, her interactions are perspectivally stable, so what is on the left will stay on the left. As soon as the infant begins to move autonomously, she no longer can rely on the egocentric coordinates and thus needs to shift to an allocentric perspective as a consequence of her own motion. An allocentric perspective is hence a consequence of self-produced locomotive activity. Interacting with objects leads to forming an even more detached, objective representation of these objects: according to Piaget, the first sensory representations of objects represent them as tied to allocation, only with locomotion and interaction the infants realize that objects can be in many different locations.

Assuming that the development of some cognitive skills is a consequence of the development of action skills makes sense if one acknowledges the idea of sensorimotor contingencies (Noë & O'Reagan 2001; see ch. 8). Accordingly, the more the infant learns to coordinate her movements and to manipulate her environment, the feedback from the environment becomes increasingly differentiated and complex. Every new interaction provides the infant with new feedback and thus the possibility of learning new sensorimotor contingencies, even if the feedback results from accidental movements or collisions. This forms the basis for developing a refined sense of objects, materials, a grasp on causality and the various object-properties in general, which will be generalized and attributed to other contexts. This reasoning is furthermore supported by the theoretical framework of neuroconstructivism (cf. Johnson and Karmiloff-Smith 2004), which claims that functional activity is a major contribution to the formation, construction and development of important structures in the nervous system.

9.3.3 Analyzing Object Concepts in Action-Frames

Another way of illustrating that possible action can be the basis of object concept development is by applying the basic insights of 'frame theory' (Barsalou 1992, 1999; Petersen 2007; Vosgerau et al. 2015). Frame theory, according to Barsalou (1992; 1999), states that frames are the general format of knowledge representation in cognition. The frame-format of representation is defined by an attribute-value structure, where the value itself can be a complex frame. Take the example of a lolly frame, as displayed in figure 1[33]: The frame for a lolly can have attributes such as 'body' and 'stick', whose values (body; stick) are complex frame in themselves:

[33] Frame graphs are a method of illustrating the structure and contents of frames. Barsalou (1992) presented simple frame graphs that have been elaborated by Petersen (2007) to a sophisticated system of illustrating frame representations. The double-encircled node represents the central node, which is what the frame stands for – in this example the frame is about lollies. The arcs represent the attributes and the individual nodes represent the values the assigned by the attribute. For example, this frame would represent that lollies have 'stick' as a general attribute, which can have the value 'stick' (as lollies are defined as having sticks) with the value 'stick' being a frame itself: It has further attributes, such as having a 'shape', which can be 'long' (cf. Petersen 2007; Petersen & Oswald 2012).

the value ‘body’ can have the attributes ‘color’, ‘shape’, ‘producer’ etc. and ‘stick’ can have the attributes ‘producer’, ‘color’ and ‘shape’ too. Possible values could then be ‘factory’, ‘green’, or ‘long’. Frame theory, especially in its refined version departing from Barsalou’s (1992) original proposal, defines attributes as being the general properties or aspects that describe a concept and the attribute-value structure being recursive. In frames, attributes assign unique values, which entails that each attribute can have only one value (out of a range of possible values). An actual lolly stick can only have one color and one shape, relative to the possible colors and shapes it can have. Being recursive means that every attribute can be a frame in itself, and thus frames are part of a larger network of frames. The frame for stick can be part of other frames as an attribute value and can be further described by attributes that have frames as values (Petersen 2007). Theoretically, there does not have to be an endpoint for frames, being part of a network of frames, though in some concepts, sensorimotor values (the actual color ‘green’ that is represented by the sensory system) can be end-values that can be no further specified (Alex Tillas, personal communication, February 2015; see also Vosgerau et al. 2015).

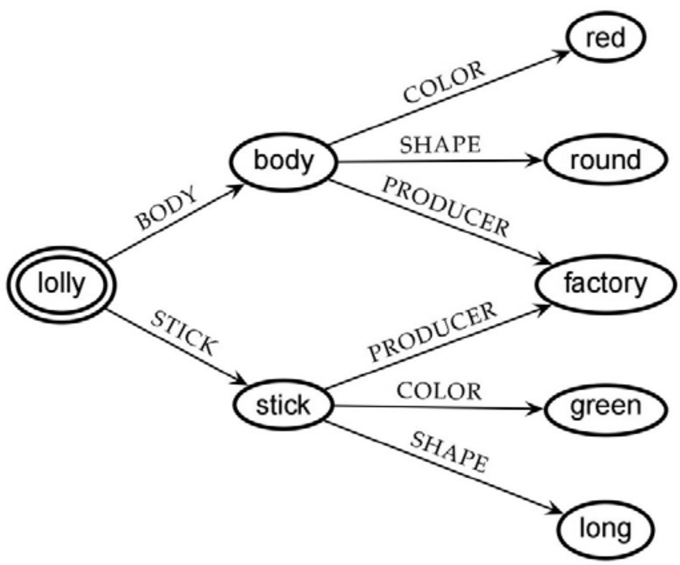

Figure 1: Lolly frame (Petersen 2007, 154)

Frame theory can account for many phenomena in linguistics, meta-science and cognition (for an overview, see Löbner 2015). Furthermore, frame analysis can illustrate and elucidate some of the claims of 'grounded cognition' theorists. For instance, frames can account for perceptually-based object representation by having sensory values as end nodes which can be not be analyzed further. In the above example, the color of the lolly's body is a part of an instantiated frame. Instantiated frames are tokened representations that actually occur in a cognitive process (such as a thought), whereas frames for type concepts (e.g. the type concept 'lolly') do not have specific values assigned by all the attributes. It might even be the case that a type concept does not represent the whole attribute structure all the time, but only when tokened. This flexibility and dynamicity allows frames to account for representing different aspects of a concept, comparable to modes of thinking or Fregean senses (Frege 1892), where a concept is represented in one aspect of many aspects possible. Thinking of dogs can represent them on one occasion as man's best friend, on the other as a threat to small children and in the next occasion dogs are thought of as smelly and furry animals that need constant attention – these thoughts involve the same type concept of 'dog', but represent different aspects of 'dog'.

To return to the lolly-example, the attribute-value 'green' in the tokened lolly concept corresponds to visual processing or retrieval of sensory information provided by visual perception. The concept 'lolly' integrates perceptual aspects which are reflected in the sensory end note and can thus be interpreted as being 'grounded', at least partly, in the sensory representation of this specific token of green. The sensory representation of this green token is the actual stored visual experience when first encountering the very object. This is what Barsalou (1999) would call a perceptual symbol that is used in simulating or reactivating a previous experience in various cognitive processes (e.g. thinking about the dog you had as a child).

In an analogous way, frames can account for motoric grounding by integrating values that cannot be further analyzed and are motoric in nature. Motoric values can be described as attribute-values that represent specific motor-parameters, which are actually used for movement control. Consider the example of the desert ant introduced in chapter 4. The desert ant represents the location of its nest in terms of its angle to the sun (and in which direction is has to turn) and the amount of steps it has to take. The ant can be justifiably said to represent the location of its nest in motoric terms, though non-conceptual: it does not assign a property (located at x) to an object (the nest), but represents it implicitly by the movements it has to execute. Figure 2 shows how this is implemented in a frame graph. The specific values in the 'heading' and 'distance' nodes are not mere numbers, but represent actual motor commands that will control the actual movements of the ant. The representation is non-conceptual, as is highlighted by the missing central node – the property of being located somewhere is not ascribed to any entity, which would not make sense anyway as this is a dynamic representation of the location of the nest that constantly changes with every change in the ant's location

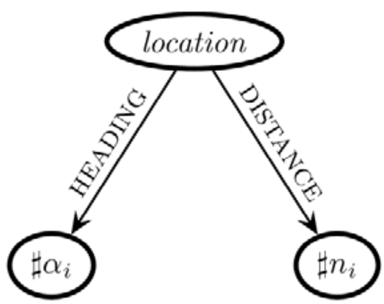

Figure 2: Frame for an ant's representation of the location of the nest (Vosgerau et. al 2015, 298)

Figure 3 shows a hypothetical frame for a conceptual representation of the location of the nest: here we have a property ascribed to an object and meet the minimum requirements of conceptual representations.

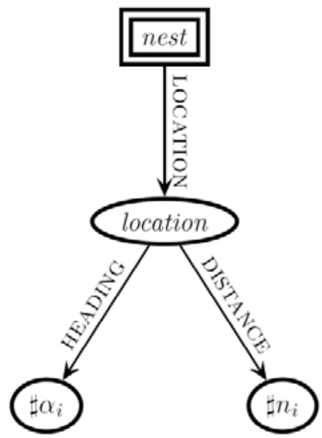

Figure 3: Frame for a conceptual representation of the location of the nest (Vosgerau et. al 2015, 299)

The frame depicted in figure 3 illustrates how a conceptual representation that is grounded in motoric values, i.e., possible movements, could look like. Objects that have crucial motoric parts are grounded in this sense. Of course, one has to accept the further premise that the motoric values cannot be further specified and are thus not frames themselves with an attribute value structure, but are the end point in analysis – the motor command that will be issued when this concept is tokened and without which the concept would not represent what it represents. In the ant example, the representation would not be about the location of the nest if the values were completely different and would thus not lead to the nest at all. Motor values are basic level values and are therefore candidates for grounding – if they were values of a higher level, they could not be the endpoint for grounding. Generally, it can be said that the idea of grounded cognition presupposes a basic level which cannot be further specified, otherwise the notion of grounding would make little sense.

The basic level grounded representation described by figure 2 becomes conceptual when further attributes and nodes are added. This happens over time in the development of organisms that are capable of developing more complex representations. The representation is conceptual, as soon as the representation assigns attributes to an object that can be analyzed independently and fulfill the other constraints on conceptual representation mentioned by Newen & Bartels (2007). Take the example of the causal indexical representation of 'within reach' again. In a frame, this would look similar to the frame in figure 2. The end values could be 'length of reaching' and 'angle of arm'. With these two values, distances in the within reach region are specified for the subject, similar to a two-place vector. To become a conceptual representation of distance for the infant, she has to learn that objects can be within or without reach for herself and for others. The crucial step in development is transcending the purely egocentric perspective and taking a more detached view of the world, realizing that all objects are spatially related to each other, and not only related to the subject, as center of all relations. From this, a more general concept of distance could be developed on the grounds of the causal indexical frame of 'within reach'. The value that is added could be 'an arm's length', which would itself be grounded in the arm length of the subject, but be generalized across subjects: a thing that is within reach for someone by being an arm's length away from that subject's arm length. The actual movement would no longer be represented, but be preserved in the value that specifies the arm length for the representing subject – always according to the ever-changing, dynamic body schema (see ch. 8.2).

Another example for an object frame with motor values is shown in figure 4. This frame graph illustrates a frame for the action of 'cutting'. In figure 4, the concept of 'cutting' is specified by attributes specifying the actual movements that have to be produced when cutting something. The movement is specified in motor parameters that would control the actual movement commands in the case of executed action. The concept of 'cutting' is thus grounded in specific movements or motor parameters. According to the simulation hypothesis (cf. Barsalou 1999; Borghi 2004; Jirak et al. 2010), every time a subject perceives or thinks about a cutting-action, the same motor patterns the subject would use for actual cutting are activated.

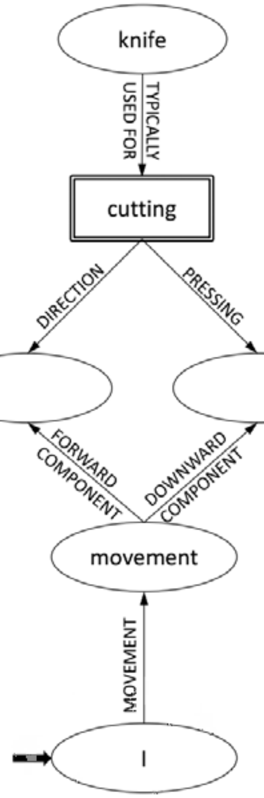

Figure 4: Frame for the concept of ‘cutting’

Figure 5 shows a frame of a specific reaching movement with the additional attribute ‘execution’, which represents the action possibility – an action, specified by motor parameters specifying a motor command can be executed or not, where the rest of the representational structure and the content of the frame representation of the concept remains the same in both conditions.

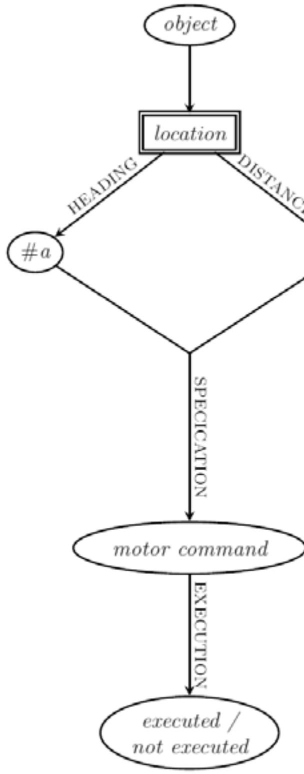

Figure 5: Frame for a specific reaching movement
(Vosgerau et al. 2015, 301)

Figure 5 shows a representation of a reaching movement for an object at a certain location. The upper part of the frame specifies location of an object and the lower part specifies the reaching movement, taken together the whole frame represents the location of an object in terms of a reaching movement. The frame in figure 4 is a more explicit representation of an object in terms of an explicitly represented subject. The lower node containing the value 'I' denotes the subject of the cutting action – the frame is a representation of the action of cutting for a specific subject. The frame also represents the typical object of cutting actions, namely of: a knife. Of course, knowing what cutting is, a subject can think of various possible objects for cutting, such as a piece of broken glass or a sharp stone. However, if the subject learns that knives are the typical objects to cut with, the representation can easily turn into a conceptual representation of knives too: in a frame graph, this would be illustrated by changing the top round node to a square double-lined one, denoting the central node.

What is more important is the idea that roughly the same representational structure can represent both an action or an object which is partly defined by the typical action associated with it, simply by shifting the focus. Obviously, a more complete representation of the concept 'knife' will have to integrate more attributes and information, such as having a handle and a blade, being made of a certain material etc. A simple 'knife' representation though could consist in terms of possible actions, having a content such as 'the object normally used for cutting'. This representation of 'knife' is just one example for a conceptual object representation in terms of possible actions, and frame analysis can help illustrate how representations integrating different aspects and even motor information could be structured.

9.4 Development of Non-Object Concepts Based on Action-Related Representation

The theory presented for cognitive abstraction on the basis of action-related representation so far only considers the development of object concepts. Object concepts are representations combining perceptual aspects and possible actions. The evidence presented supports the claim that object representation involves - to a varying degree – the representations of possible actions in terms of possible movements. Any theory of grounded cognition should also be able to account for the development more abstract concepts. There have been promising attempts to do so, namely by Barsalou (1999), who describes abstract concepts such as 'truth' and 'negation' on the basis of matching expected representations. Prinz (2005) accounts for abstract concepts such as 'democracy' by stored sensory knowledge (the house of parliament; the act of voting; a chancellor giving a speech) together with semantic information compiling the meaning of abstract concepts. The theory of action-related representation developed here is, at least, compatible with these existing approaches dealing with abstract concepts. In addition, action-related representation provides a new angle for abstract cognition in terms of focusing primarily on *abstraction mechanisms* rather than on abstract concepts. The core idea is that abstraction mechanisms are at work from the very beginning of development, from the moment a subject starts to interact with the environment. Thus, the main aim of the theory presented is not to account for the possession of individual concepts that are considered to be abstract, but rather how abstract cognitive processes in general could be described and explained on a basic level. The accounts of neo-empiricists like Barsalou (1999) or Prinz (2005) have a strong focus on perceptual representation as possible vehicles of grounding, whereas the action-related accounts discussed and developed here shift attention towards action generation and output while recognizing the importance of perceptual input.

An action-related approach to concept development exceeding object concepts will make the following claims:

- The basis of all abstraction is generalization, which is involved in processes such as the transition from purely egocentric to allocentric frame of reference.
- Abstraction is also present in representations becoming detached from the specific situations in which they were first formed or deployed.
- The abilities required for forming object concepts and for transferring this knowledge involve abstraction mechanisms and these mechanisms are at the basis of all other abstract cognitive operations.

Thinking of abstraction this way, concept development as a whole is an abstract process and can, in principle, be traced back to sensorimotor representations. As a consequence, the traditional dichotomy of abstract vs. concrete concepts or thinking is diluted by the action-related approach, if not given up completely. This implies that the debate about concept development should no longer focus simply on the distinction of “concrete” concepts that refer to actual objects in the world and “abstract” concepts that denote 'things one cannot touch'. Rather, abstract cognition should be conceived of as a process or a cognitive operation in play whose output involves concepts that no longer refer to concrete particulars, and any theory of abstraction should be concerned to a greater extend with explaining the cognitive operations at work in abstraction, than with the meaning of thoughts involving abstract concepts.

A vast body of existing research approaches abstract concepts from different perspectives. For example, take the work of Lakoff & Johnson (1980; 1999) who claim that the meaning of abstract concepts is provided by conceptual metaphor. The abstract meaning of 'love' is thus explicated by concrete bodily and physical states and activity, such as in the metaphor 'love is a journey', which refers to the concrete efforts a traveler has to make (cf. Lakoff & Johnson 1999). They claim that we understand abstract concepts by using concrete metaphors that often refer to physical action, thus the abstract concepts always have to be "translated" into the concrete realm to be cognitively accessible. This rough sketch of their theory of conceptual metaphors and their role in cognition is one example for an approach of how to understand abstract concepts in cognition.

Another prominent example is that offered by Prinz (2005), who accounts for abstract concepts by linguistic labels that are part of an inferential network of labels. These other labels ultimately break down into perceptual components of either sounds or gestures or objects involved, actions and social situations (as in the case of 'democracy'). In addition, there are approaches accounting for numerical cognition in terms of embodiment (cf. Dehaene et al. 1993; Fischer & Brugger 2011; Tschentscher et al. 2011; Fischer 2012), which provide evidence for a functional link between finger counting and number processing.

The action-related approach is aims at a slightly different solution. Take the concept 'cup'. It is, quite plausibly derived from encounters with actual cups. The general concept 'cup' is an abstraction from the specific features of cups, and means something like 'cuphood'. Central for the 'cup' concept is the way cups allow for interaction, the interactive properties of cups. Cups have handles, they can be grasped, they are for drinking and they are containers for all drinkable liquids, from water and coffee to wine. The concept of a container related to, or being part of the concept 'cup' is more abstract: it refers to all kind of objects that can be used for storing liquids and other material. There is still an action-related component in 'container', e.g., that containers can be used for transporting liquids from A to B, that hands taken together can form a container, etc. Nevertheless, container as such is a rather general concept not referring to specific objects anymore, but obviously still connected to specific objects via the functional aspects that play a role in interaction. Thus, at least some aspects of the concept 'container' are, probably in a more indirect way, grounded in possible actions. The background cognitive operations allowing for the abstraction are generalization and classification on the basis of interaction encounters.

9.4.1 Abstraction Mechanisms for Classification

What has been described so far is what level of abstraction occurs at which stage of development and how it correlates with the development of other cognitive abilities, such as object permanence, mindreading (i.e., the ability of ascribing mental states to other subjects) or perspective taking (see ch. 9.2.2). What is still missing is a description of how the actual abstraction mechanism at work can be described. In discussing the development of general ideas (general concepts) in the light of Locke's problem of circularity (as raised e.g. by Berkeley 1710/1957), Tillas (2014) proposes a model of cognitive abstraction and provides empirical evidence supporting the model. According to this model, raw data, i.e., perceptual information during an encounter with an individual object, is stored in long-term memory. The representations formed this way are scanned by a scanning process checking for matching features. Similar representations will be stored in a similar location. These bundles of representations initiate the abstraction process. Two conditions have to obtain for the initiation of the abstraction process: a sufficient number of stored representations in one location and the stored representations showing a sufficient degree of diversity. The output of the abstraction process then is a representation of a category, such as the category tree that is based on an abstraction over all previous tree encounters. The abstract representation is more general, in that it picks out all members of a given category and is able to represent information that has not been perceived on previous encounters. The underlying mechanism is a process analogous to Hebbian learning: co-occurring representations builds stronger connections and are thus more accessible, while representations that are less connected to another are harder to access (Hebb 1949). Tillas presents a vast body of evidence, arguing mainly that early visual processing and the way the sensory system is constructed can account for most of the claims. Especially pattern recognition abilities are able to grant the detection of similarities of class members and thus the adequate storage in the right location on which the abstraction process could work. Pattern recognition abilities can explain the detection of similarities without presupposing a notion or a concept of similarity. Thus, a fundamental similarity detector is assumed to be hardwired deep in the perceptual systems and the cognitive operations emerging from them. The circularity worries therefore do not arise as they did for Locke, as the respective similarity is already given on a pre-conceptual, pre-categorical level. Similarity allows for categorizing, but the mechanism is so low-level in nature that no questionable presuppositions have to be made (for the detailed account and the discussion of the empirical evidence, see Tillas 2014).

This should not be the place to discuss whether Tillas' account actually solves the circularity problem while presenting a viable explanation for cognitive abstraction. Problems might arise on the level of similarity detection already in early vision, this shifting the problem – the similarity presupposed might no longer be a concept of similarity, but still similarity has to be detectable before categories can be formed. From this it would follow that a category is what is detected by the similarity detection mechanisms. If one allows for certain inbuilt systematic structures in our cognitive organization, then this would only be a problem for a hardcore empiricist, who would have to show how these operations are formed without presupposing inbuilt, hardwired similarity detectors. I leave the question open if this can be done easily or at all, recognizing that the strength of Tillas' account lies in it building on very low-level cognitive structures and thus gaining plausibility and explanatory potential from a developmental and evolutionary point of view.

Applied to action-related representation, at the basic level, arbitrary movements and simple interactions with objects create feedback that is represented, thus stored in memory. Similar feedback on other occasions is stored at the same location and thus able to initiate a scanning process. What needs to be established are links and connections between sensory and motor representation, forming sensorimotor representation, which contain perceptual (sensory input, proprioceptive input etc.) and motor information (motor parameters, motor commands etc.). The basic action-related representations thus contain information from different channels: sensory input plus motor output information is linked and processed. The scanning process will detect the common features of newly acquired and stored representations and will allow for storage in appropriate loci. Once the threshold and quality of representation in one locus is reached, the abstraction process is initiated, forming a category representation as output. This category will be the precursor of an action-related object concept, such as: 'is a thing I can grasp, implying this and that movement and having certain perceptual features'. A basic category comes into existence only when repeated successful interactions occurred and the movements involved become systematic. This allows for faster detection of action possibilities and action selection in further encounters with objects of a given kind. After learning the interaction possibilities of, say, tennis balls, the experimental phase where the subject will find out by random trial and error interaction will be shorter and goal directed interaction can happen more quickly, as the object is categorized along interaction dimension. To illustrate this, a child might learn that a tennis ball will roll if pushed, will bounce back if thrown against walls and will float on water, and of course how to identify them visually or by texture. This allows in a first step for simple classification, which involves discriminating the object based on similarity. At later stages this knowledge will be used to form a category which is the basis for the concept 'tennis ball' – the tennis ball will be represented as tennis ball, implying disparate information, exceeding the mere interactional aspects for the subject. Before a general 'ball' category can be formed, the subject needs to have similar interaction experience with other objects, which means that from similar motor output, a similar feedback is provided, stored and compared with the existing representations. If the other representations from objects are sufficiently similar, the subject will be able to form a general 'ball' category – all object that behave in similar ways when interacting with them and thus having similar interactive properties – with similarity in appearance also being given. The appearance thus is not sufficient for an action-related category, at least not on the basic level: if the subject encounters a heavy concrete sculpture that looks like a football but is too heavy to interact with in a ball-like fashion, this object will most likely not represented by or contribute to forming the 'ball' category, as the crucial action-related representation is not given. Later in development, where more abstract concepts are available, this object can be identified as an abstracted version of a ball, but the initial ball category will pick out objects that can be interacted with according to stored action-related information. This way, the action-related information will always be a part of the category but will not be necessarily activated in later stages of development when a complex net of (linguistic) concepts has been formed and other aspects of objects might be primarily represented when using a concept in a given situation, such as 'this year's Soccer World Cup ball has been entirely developed in Herzogenaurach'.

Jung and Newen (2011) present an approach to distinguishing different types of knowledge formats. Representative for these formats are different representation types, which are identifies as propositional, image-based and sensorimotor (cf. Jung & Newen, 2011, 96). This reflects the way of concept development I have just presented: Before a full-blown conceptual representation can be developed, sensorimotor information and perceptual information form the crucial representational contents, before this knowledge is transformed into a propositional format, which typically presupposes the possession of linguistic concepts.

9.5 Summary

Thus, the account of concept development presented can be explained with abstraction mechanisms found in general category formation accounts, such as Tillas (2014). This is not the place for an extensive discussion of all other possible abstraction accounts, all that should be demonstrated is that abstraction and concept development is in line with contemporary research on conceptual development in animals and humans and enjoys thus, alongside philosophical plausibility, a strong empirical foundation.

Action-related information is part of actual object concepts which are formed (to a certain degree) on the basis of action-related representation. Other, less obviously action-related concepts can also emerge from concepts that have a more significant action relation, such as shown in the example of 'container' as a rather abstract concept that might well be derived on the basis of interactions with various cups and glasses or action involving one's hands as cups. Drawing boundaries and lines of demarcation in the development of concepts according to degrees of abstraction would be impossible – the gradual transition model presented here should illustrate this fact about cognitive development. On all stages, perceptual information will be combined and processed together with action-related information, either through one's own actions or observing others acting. Being able to interact successfully with one's world allows for a better understanding of other subjects' actions and at the same time observing others facilitates the development of behavioral competencies. The cognitive operations thus described are all mutually interdependent and are initiated from the very onset of cognitive development. What cannot be denied is the great role action plays for cognitive development in general in all important aspects: successful goal realization, understanding objects and developing systematic categories, representing the world in terms of interaction possibilities and thus, increasingly, enhancing the performance of subjects in novel situations – all these phenomena can be accounted for by action-related representations being systematically integrated in the cognitive system.

10 Conclusion: Grounding Cognition in Action?

The discussion of action-related representations and the general account developed in the previous chapters is a contribution to a more general theory of grounded cognition. Action-related representations are a plausible option for cognitively foundational representations, precisely because action-related representation is the most basic kind representation that exists across a wide variety of species, from insects to humans, as their original function is given in terms of behavior-guidance. Action-guiding mechanisms are crucial for the individual's continued existence – living organisms need to make an effort to maintain the system's stability.

However, action-related representation is not only describing a kind of representation that crucially guides the behavior of animals, but is equally an account of the role action-related representations have for cognitive abilities that are indirectly related to action-guidance. This implies that cognitive abilities such as classification of environmental features and objects can be explained on the basis of action-related representations and the interactions they guide. Being basic representations, action-related representations give rise and are thus the grounds for other kinds of representations. Moreover, abstraction mechanisms can be identified in the gradual development of increasingly complex action-related representations, such as perspectival shift and explicit property representation. These two aspects are indications of more abstract, more sophisticated ways of representing features of a subject's environment, thus providing the subject with a greater behavioral flexibility. With the ability to represent and attend to different aspects of an object, more possibilities to interact with this object come into existence. This in turn is evolutionary advantageous, as it allows animals better ways to adapt and adjust their behavior in accordance with changes of the environment and the animals body.

Action-related representation can explain the development of classification processes that are the basis for categorization and concept development. Central to this development is abstracting from a purely egocentric, essentially causal indexical way of representing action possibilities, which has been identified as the basic feature common to basic action-related representations for animals and humans alike (see ch. 4; 8). Accordingly, action-related abstraction can be described as developing the ability to take a new perspective, which becomes manifest in the gradual transition from purely, implicitly egocentric to action-related representations involving other subjects. Without this step in development, the 'primordial subject-object' entanglement (Piaget 1977) will never be transcended and representations that explicitly distinguish subject and object are not available for the cognitive system. Whereas many creatures that use basic action-related representation to guide their behavior will always remain on the level of pure egocentricity, many species, including primates and humans are able to develop a detached perspective and thus explicitly represent objects in their independent existence, as well as other subjects as cognitive agents with intentional mental states.

As action-related representations are representations that essentially involve (possible) movements of subjects, they are fundamentally referring to bodily aspects of the respective subjects. Cognitive abilities that can be described as grounded in action-related representations are thus grounded in representations of possible movements and are able to provide a profound meaning of the claim that cognitive abilities are grounded in sensorimotor representations. Cognitive abilities, such as classification and, at later stages categorization develop on the grounds of concrete interactions with the subject's environment. Aspects of the environment are given a 'motoric meaning', based on the action possibilities that a subject is able to represent. Features of objects are represented in the most basic way: As possible movements corresponding to these features, such as a grasping movement, which involves a certain grip aperture, represents the width and distance of an object in egocentric coordinates. Following Bickhard (2002), an object representation can be developed on the basis of grouping related interaction possibilities an object allows for or corresponds to. This probably what Gibson (1986) had in mind, when speaking about perceiving an object solely in terms of the object's affordances. Thus, the crucial step in developing more abstract representations is the transition from simply representing features in terms of movements, to grouping represented features by means of corresponding represented action possibilities and thereby developing an object representation. Once an object is constituted via action-related representation, this knowledge can be used for further classification of objects on that basis.[34]

The representation of a class of objects can thus be described as grounded in possible action. Action-related representations provides a new angle to the problem of grounding cognitive abilities, in that the focus is on interaction in accordance to features of the environment instead of mere visual perception of object's features and a classification and conceptualization on this perceptual basis (cf. Barsalou 1999). In this sense, the account of action-related representation is complementary to other accounts of grounded cognition, such that the meaning of sensorimotor representations for cognition can be given in terms of visual processing for action guidance, where visual features are in the first instance represented in an action format too. At later stages of development and further complexity of the cognitive system, purely visual processing is possible without direct implications for action any longer, but the arguments and evidence discussed in the previous chapters make it very plausible to assume that vision at early (ontogenetic and phylogenetic) stages subserves action guidance and is thus part of action-related representation.

Another aspect that can be understood as grounded in action is the development of a 'self'-concept. As could be shown, the self-relation is essentially implied in even the most basic action-related representations (ch. 4). This self-aspect that is an implicit part of the basic action-related representation does not yet enable the subject to think about herself as herself. On the basic level, this can be understood in terms of a 'feeling of

34 Of course, these abstraction processes required further cognitive structures and abilities, such as memorizing and association, together with basic pattern recognition. This is no longer the scope of action-related representation, thus the existence of these structures in many species is simply presupposed.

agency' and (Synofzik et al. 2008a, 2008b), which is mainly based on proprioceptive and sensory feedback and non-compositional, thus does not have a property object structure. The feeling of agency is a product of interaction and the associated perceptual feedback and thus a natural product of a subject's constitution as an active being. This early notion of a sense of agency is clearly grounded in agency and an essential aspect of action-related representation. Explicit self-representations, which are the basis for judgments of agency, are more complex representations, having a property-object structure, and being compositional. The development of these higher order agency representations can be explained by the development of increasingly complex action skills, which also generates increasingly complex proprioceptive and sensory feedback. Self-produced movement, such as toddling around generates a different sensory feedback input, while social interactions with other agents create 'social feedback'. The child learns that objects are persistent entities that feature in observed actions of other subject's. This way, the child might learn about action types that can be performed other subjects too, and with this knowledge, the other subjects can be represented as agents. Once this is established, the child is able to represent possible actions for other subjects, such as reaching for an object that is out of reach for the child but can be reached by an adult. Representing another subject's action possibilities is the precursor for developing a concept of 'self', as it implies an broader notion of agency that distinguishes between own actions and other subjects' actions. Once this distinction can be made by a child, the merely implicit representation of the child's own agency will become a richer notion of selfhood, which can enter representations as an explicitly represented aspect. These transitions reflect abstraction processes on the basis of action-related representations and show how a self-concept arises out of developing a more complex behavioral repertoire that for this important ability of perspective taking.

Not all aspects of abstract cognition can be easily explained within this framework. Mathematical cognition and complex symbolic thought are of such a high degree of abstraction, which is why it is currently impossible for any theory of abstraction to do more than suggesting some foundational cognitive operations that can be grounded in low level representations. However, it is widely accepted that linguistic abilities crucially involved in enabling and shaping these higher-order cognitive operations (cf. Rakoczy 2010). Higher-order thought is not conceivable without a systematic and compositional language to express and further determine thoughts. However, the gap seems no longer to impossible to close, and the various approaches to embodied language are offering valuable accounts of how language and low-level cognitive processes are interdependent or related (Glenberg & Kaschak 2002; Pulvermüller 2005).

The discussion and analysis of action-related representation is a contribution to closing the gap between basic-level and higher-order cognitive operations. Action-related representations are fundamental for goal-oriented, successful interactions with one's environment. At the same time, they are the basis for simple classification and discrimination in terms of action-possibilities which in turn is the foundation for cognitive abilities involving generalization and conceptual representation. On an action-related basis, the development of object concepts and action concepts can be explained and be related to the motoric processes that generate and guide movements. If the development of linguistic communication is understood as extending one's behavioral repertoire by means of symbolic, representational communication that facilities cognitive processes and is crucial for reaching complex goals, on both an individual and social level, the foundations for language development are quite plausibly located within an action-related framework too.

This last claim needs further elaboration though, and future research crucially has to integrate the different philosophical, psychological and neurobiological perspectives that have been presented in the previous chapters. Action-related representations are central to the development of cognitive abilities, bringing together perception, action and higher-order cognition across different species of basically all levels of cognitive complexity. The general account of action-related representation developed here is supposed to offer an applicable definition of action-related cognitive processes to empirical and philosophical research, by uniting the central elements of basic action cognition in one account: implicit self-related representation of environmental features in terms of possible movements of the agent. All other aspects of action-related cognition are grounded in these essential elements, and are the results of gradual transitions and variations of the core aspects.

Bibliography

Agre, P. (1997), *Computation and Human Experience*, Cambridge University Press, Cambridge/New York.

Allen, J. W. P. & Bickhard, M. H. (2013), 'Stepping off the pendulum: Why only an action-based approach can transcend the nativist-empiricist debate', *Cognitive Development* 28, 96–133.

Anderson, D. I.; Campos, J. J.; Witherington, D. C.; Dahl, A.; Rivera, M.; He, M.; Uchiyama, I. & Barbu-Roth, M. (2013), 'The role of locomotion in psychological development', *Frontiers in Psychology* 4(440).

Anderson, M. L. (2003), 'Embodied Cognition: A field guide ', *Artificial Intelligence* 149(1), 91–130.

Anderson, M. L. & Rosenberg, G. H. (2008), 'Content and action: The guidance theory of representation', *Journal of Mind and Behavior* 29(1–2), 55–86.

Auersperg, A. M. I.; von Bayern, A. M. P.; Gajdon, G. K.; Huber, L. & Kacelnik, A. (2011), 'Flexibility in Problem Solving and Tool Use of Kea and New Caledonian Crows in a Multi Access Box Paradigm', PLoS ONE 6(6), e20231.

Barsalou, L. (1999), 'Perceptual symbol systems', *Behavioral and Brain Sciences* 22, 577–609.

Barsalou, L. W. (1992), Frames, Concepts and Conceptual Fields, *in* A. Lehrer & E. Feder Kittay, ed., 'Frames, Fields and Contrasts', Erlbaum, Hillsdale, Hove and London, 21–74.

Barsalou, L. W. (2008), 'Grounded Cognition', *Annual Review of Psychology* 59, 617–645.

Barsalou, L. W.; Simmons, W. K.; Barbey, A. K. & Wilson, C. D. (2003), 'Grounding conceptual knowledge in modality-specific systems', *Trends in Cognitive Sciences* 7(2), 84–91.

Beauchamp, M. & Martin, A. (2007), 'Grounding object concepts in perception and action: evidence from fMRI studies of tools', *Cortex* 43, 461–468.

Berkeley, G. (1710/1957), A treatise concerning the principles of human knowledge, Bobbs-Merrill, Indianapolis, IN.

Berti, A. & Frassinetti, F. (2000), 'When far becomes near: Remapping of space by tool use', *Journal of Cognitive Neuroscience* 12(3), 415–420.

Bickhard, M. (2002), The biological emergence of representation, *in* L. Smith T. Brown, ed., 'Emergence and Reduction: Proceedings of the 29th Annual Symposium of the Jean Piaget Society', 105–131.

Bickhard, M. (2009), 'Interactivism: A Manifesto', *New Ideas in Psychology* 27(1), 85–95.

Bickhard, M. (2009), 'The Interactivist Model', *Synthese* 166, 547–591.

Bickhard, M. H. (2003), 'Some Notes on Internal and External Relations and Representation', *Consciousness and Emotion* 4(1), 101–110.

Bickhard, M. H. (1999), 'Interaction and representation', *Theory & Psychology* 9(4), 435–458.

Borghi, A. M. (2004), 'Object concepts and action: Extracting affordances from objects parts 115 (1), 69-96', *Acta Psychologica* 115(1), 69–96.

Bosbach, S.; Cole, J.; Prinz, W. & Knoblich, G. (2005), 'Inferring another's expectation from action: The role of peripheral sensation.' *Nature Neuroscience* 8, 1295–1297.

Brentano, F. (1874), Psychologie vom empirischen Standpunkte, Duncker & Humblot, Leipzig.

Brooks, R. (1991), 'Intelligence Without Representation', *Artificial Intelligence* 47, 139–159.

Campbell, J. (1993), The role of physical objects in spatial thinking, *in* N. Eilan; R. McCarthy & B. Brewer, ed., 'Spatial Representations: Problems in Philosophy and Psychology', Blackwell, Oxford, Cambridge MA, 65–95.

Campbell, J. (1994), Past, Space, and Self, MIT Press.

Cardinali, L. (2011), Body schema plasticity after tool-use, Université Claude Bernard – Lyon I.

Cardinali, L.; Brozzoli, C. & Farne, A. (2009), 'Peripersonal Space and Body Schema: Two Labels for the Same Concept?', *Brain Topography* 21, 252–260.

Cardinali, L.; Jacobs, S.; Brozzoli, C.; Frassinetti, F.; Roy, A. C. & Farne, A. (2012), 'Grab an object with a tool and change your body: tool-use-dependent changes of body representation for action', *Experimental Brain Research* 218, 259–271.

Carlson, T. A.; Alvarez, G.; an Wu, D. & Verstraten, F. A. (2010), 'Rapid Assimilation of External Objects Into the Body Schema', *Psychological Science* 21(7), 1000–1005.

Carman, T. (1999), 'The Body in Husserl and Merleau-Ponty', *Philosophical Topics* 27(2).

Chaminade, T.; Meltzoff, A. & Decety, J. (2005), 'An fMRI study of imitation: action representation and body schema', *Neuropsychologia* 43(1), 115–127.

Chao, L. L. & Martin, A. (2000), 'Representation of Manipulable Man-Made Objects in the Dorsal Stream ', *NeuroImage* 12(4), 478-484.

Chemero, A. (2011), *Radical Embodied Cognitive Science*, A Bradford Book.

Chemero, A. (2003), 'An Outline of a Theory of Affordances', *Ecological Psychology* 15(2), 181–195.

Chomsky, N. (1959), 'Review of Skinner's Verbal Behavior', *Language* 35, 26–38.

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

Clark, M. A.; Merians, A. S.; Kothari, A.; Poizner, H.; Macauley, B.; Rothi, L. J. G. & Heilman, K. H. (1994), 'Spatial planning deficits in limb apraxia', *Brain* 117, 1093–106.

Colby, C. L. (1998), 'Action-Oriented Spatial Reference Frames in Cortex', *Neuron* 20, 15–24.

Craighero, L.; Fadiga, L.; Rizzolatti, G. & Umilta, C. (1999), 'Action for Perception: A Motor-Visual Attentional Effect', *Journal of Experimental Psychology* 25(6), 1673–1692.

Davidson, D. (1967), 'Truth and meaning', *Synthese* 17, 304–323.

Dehaene, S.; Bossini, S. & Giraux, P. (1993), 'The mental representation of parity and number magnitude.' *Journal of Experimental Psychology* 122(3), 371–396.

Dienes, Z. & Perner, J. (1999), 'A Theory of Implicit and Explicit Knowledge', *Behavioral and Brain Sciences* 22(5), 735–755.

Dretske, F. (1986), Misrepresentation, *in* R. Bogdan, ed., 'Belief: Form, Content and Function', Clarendon Press, Oxford, 17–36.

Dretske, F. (1981), *Knowledge and the Flow of Information*, MIT Press, Cambridge, MA.

Dumontheil, I. (2014), 'Development of abstract thinking during childhood and adolescence: The role of rostrolateral prefrontal cortex', *Developmental Cognitive Neuroscience* 10, 57–76.

Edelman, S. (2008), *Computing the Mind: How the Mind Really Works*, Oxford University Press.

Ellis, R. & Tucker, M. (2000), 'Micro-affordance: The potentiation of components of action by seen objects', *British Journal of Psychology* 91, 451–471.

Evans, G. (1985), Molyneux's Question, *in* A. Phillips, ed., 'Gareth Evans: Collected Papers', Clarendon Press, Oxford.

Evans, G. (1982), *The Varieties of Reference*, Oxford University Press, Oxford.

Fischer, M. & Brugger, P. (2011), 'When digits help digits: spatial-numerical associations point to finger counting as prime example of embodied cognition', *Frontiers in Psychology* 2, 1–7.

Fischer, M. H. (2012), 'A hierarchical view of grounded, embodied, and situated numerical cognition', *Cogn Process* 13(Suppl. 1), 162–164.

Fodor, J. (1975), *The Language of Thought*, Crowell, New York.

Fodor, J. A. (1990), Information and Representation, *in* P. P. Hanson, ed., 'Information, Language, and Cognition', University of British Columbia Press.

Fodor, J. A. (1983), *The modularity of mind: an essay on faculty psychology*, MIT Press.

Fodor, J. & Pylyshyn, Z. (1981), 'How direct is visual perception?: Some reflections on Gibson's "ecological approach"', *Cognition* 9(2), 139–196.

Frege, G. (1892), 'Über Sinn und Bedeutung', *Zeitschrift für Philosophie und philosophische Kritik* 100, 20–50.

Frege, G. (1918–1919), 'Der Gedanke. Eine Logische Untersuchung', *Beiträge zur Philosophie des deutschen Idealismus I*, 58–77.

Friston, K. J. & Price, C. J. (2001), 'Dynamic representations and generative models of brain function', *Brain Research Bulletin* 54(3), 275–285.

Gallagher, S. (2009), Body image/body schema, *in* T. Bayne; A. Cleeremans & P. Wilken, ed., 'Oxford Companion to Consciousness', Oxford University Press, Oxford, 117–119.

Gallagher, S. (1995), Body Schema and Intentionality, *in* N. Eilan; A. Marcel & J. L. Bermúdez, ed., 'The Body and the Self', MIT Press, Cambridge MA.

Gallagher, S. (2001), Dimensions of Embodiment: Body Image and Body Schema in Medical Contexts, *in* S. Kay Toombs, ed., 'Handbook of Phenomenology and Medicine', Springer Netherlands, 147–175.

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

Gallagher, S. & Cole, J. (1995), 'Body Schema and Body Image in a Deafferented Subject', *Journal of Mind and Behavior* 16, 369–390.

Gallagher, S. & Meltzoff, A. N. (1996), 'The earliest sense of self and others: Merleau-Ponty and recent developmental studies', *Philosophical Psychology* 9(2), 211–236.

Gallese, V. (2007), The Ventro-dorsal Stream: Parieto-premotor Neural Circuits and Their Role in Primate Cognition, *in* F. Mast & L. Jäncke, ed., 'Spatial Processing in Navigation, Imagery and Perception', Springer US, 329–352.

Gallese, V. & Gallese, C. V. (2007), 'The “conscious” dorsal stream: Embodied simulation and its role in space and action conscious awareness', *Psyche*, 1--20.

Gallese, V. & Sinigaglia, C. (2011), 'What is so special about embodied simulation?', *Trends in Cognitive Sciences* 15(11), 512–519.

Gallistel, C. (1993), *The Organization of Learning*, The MIT Press, Cambridge MA.

Gallup, G. G., J. (1979), *Self-recognition in chimpanzees and man: A developmental and comparative perspective*, Plenum Press, New York.

Gibson, J. J. (1986), *The Ecological Approach to Visual Perception*, Houghton Mifflin, Boston.

Gibson, J. J. (1966), *The Senses Considered as Perceptual Systems*, Houghton Mifflin, Boston.

Glasersfeld, E. von (1990), 'An exposition of constructivism: Why some like it radical', *Journal for Research In Mathematics Education* Monograph 4, 19–29, 195–210.

Glenberg, A. M. & Kaschak, M. P. (2002), 'Grounding language in action', *Psychonomic Bulletin & Review* 9, 558–565.

Goldman, A. & de Vignemont, F. (2009), 'Is social cognition embodied?', *Trends in Cognitive Science* 15(10), 154–159.

Grossman, M.; Anderson, C.; Khan, A.; Avants, B.; Elman, L. & McCluskey, L. (2008), 'Impaired action knowledge in amyotrophic lateral sclerosis', *Neurology* 71(18), 1369–1401.

Hansell, M. (2009), *Built By Animals*, Oxford University Press, New York.

Harnad, S. (1990), 'The Symbol Grounding Problem', *Philosophical Explorations* 42, 335–346.

Hatfield, G. (2002), Perception as Unconscious Inference, *in* D. Heyer & R. Mausfeld, ed., 'Perception and the Physical World: Psychological and Philosophical Issue in Perception', John Wiley & Sons, Ltd.

Hebb, D. O. (1949), *Organization of behavior: A neuropsychological theory,* Wiley, New York.

Heft, H. (1989), 'Affordances and the Body: An Intentional Analysis of Gibson's Ecological Approach to Visual Perception', *Journal for the Theory of Social Behaviour* 19(1), 1–30.

Hommel, B. (2009), 'Action control according to theory of event coding (TEC)', *Psychological Research* 73, 512–526.

Hommel, B. (2004), 'Event files: feature binding in and across perception and action', *Trends in Cognitive Sciences* 8(11), 494–500.

Hommel, B.; Müsseler, J.; Aschersleben, G. & Prinz, W. (2001), 'The Theory of Event Coding (TEC): A Framework for Perception and Action Planning', *Behavioral and Brain Sciences* 24, 849–937.

Hoshi, E. & Tanji, J. (2000), 'Integration of target and body-part information in the premotor cortex when planning action', *Nature* 408(6811), 466–470.

Huber, D.; Gutnisky, D. A.; Peron, S.; O'Connor, D. H.; Wiegert, J. S.; Tian, L.; Oertner, T. G.; Looger, L. L. & Svoboda, K. (2012), 'Multiple dynamic representations in the motor cortex during sensorimotor learning', *Nature* 484(7395), 473–478.

Hurley, S. (2001), 'Perception and Action: Alternative Views', *Synthese* 129, 3–40.

Hutto, D. D. & Myin, E. (2012), *Radicalizing Enactivism: Basic Minds Without Content*, The MIT Press.

Iachini, T.; Borghi, A. M. & Senese, V. P. (2008), 'Categorization and sensorimotor interaction with objects', *Brain and Cognition* 67(1), 31–43.

Ingle, D. (1973), 'Two Visual Systems in the Frog', *Science* 181(4104), 1053–1055.

Jacob, P. & Jeannerod, M. (2003), *Ways of Seeing*, Oxford University Press.

Jacobs, L. F. & Liman, E. R. (1991), 'Grey squirrels remember the locations of buried nuts', *Animal Behaviour* 41, 103–110.

James, T. W.; Culham, J.; Humphrey, G. K.; Milner, A. D. & Goodale, M. A. (2003), 'Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study', *Brain* 126, 2463–2475.

Jeannerod, M. (2007), Consciousness of Action, *in* M. Velmans & S. Schneider, ed., 'The Blackwell Companion to Consciousness', Blackwell, 540–550.

Jeannerod, M. (2006), *Motor Cognition*, Oxford University Press.

Jeannerod, M. (1997), *The Cognitive Neuroscience of Action*, Blackwell, Oxford.

Jeannerod, M. (1995), 'Mental Imagery in the Motor Cortex', *Neuropsychologia* 33(11), 1419–1432.

Jeannerod, M.; Arbib, M.; Rizzolatti, G. & Sakata, H. (1995), 'Grasping objects: the cortical mechanisms of visuomotor transformation', *Trends in Neurosciences* 18(7), 314–320.

Jirak, D.; Menz, M. M.; Buccino, G.; Borghi, A. M. & Binkofski, F. (2010), 'Grasping language: A short story on embodiment', *Consciousness and Cognition* 19(3), 711–720.

Johnson, M. H. & Karmiloff-Smith, A. (2004), Neuroscience Perspectives on Infant Development, *in* G. Bremner & A. Slater, ed., 'Theories of Infant Development', Blackwell, Oxford, 121–144.

Johnson-Frey, S. (2004), 'The neural bases of complex tool use in humans', *Trends in Cognitive Sciences* 8(2), 71–78.

Jung, E.-M. & Newen, A. (2011), Understanding Knowledge in a New Framework: Against Intellectualism as a Semantic Analysis and an Analysis of the Mind, *in* A. Newen; A. Bartels & E.-M. Jung, ed., 'Knowledge and Representation', Stanford, Calif.: CSLI Publications, 79–105.

Kelly, S. D. (2002), 'Merleau-Ponty on the body', *Ratio* 15(4), 376–391.

Lakoff, G. & Johnson, M. (1999), *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*, Basic Books, New York.

Lakoff, G. & Johnson, M. (1980), *Metaphors we Live By*, University of Chicago Press.

Lewis, E. & Lloyd, D. M. (2010), 'Embodied experience: A first-person investigation of the rubber hand illusion', *Phenomenology and the Cognitive Sciences* 9(3), 317–339.

Löbner, S. (2015), Functional Concepts in Frames, *in* Th. Gamerschlag; D. Gerland; R. Osswald & W. Petersen, ed., 'Meaning, Frames, and Conceptual Representation', Düsseldorf University Press, Düsseldorf, 13–42.

Longo, M. R.; Azanon, E. & Haggard, P. (2010), 'More than skin deep: body representation beyond primary somatosensory cortex', *Neuropsychologia* 48(3), 655–668.

Magnie, M.-N.; Ferreira, C.-T.; Giusiano, B. & Poncet, M. (1999), 'Category specificity in object agnosia: Preservation of sensorimotor experiences related to objects', *Neuropsychologia* 37(1), 67–74.

Mandik, P. (2005), Action-oriented perception, *in* A. Brook & K. Akins, ed., 'Cognition and the Brain: The Philosophy and Neuroscience Movement', Cambridge University Press, New York, 284–305.

Maravita, A. & Irik, A. (2004), 'Tools for the body (schema)', *Trends in Cognitive Sciences* 8(2), 79–86.

Meltzoff, A. N. & Moore, M. K. (1983), 'Newborn Infants Imitate Adult Facial Gestures', *Child Development* 54, 702–709.

Merleau-Ponty, M. (1945/2012), *Phenomenology of Perception*, Routledge, New York.

Millikan, R. (1994), Biosemantics, *in* St. Stich, ed., 'Mental Representation. A Reader', Blackwell, Cambridge MA, Oxford, 243–258.

Millikan, R. G. (1995), 'Pushmi-Pullyu Representations', *Philosophical Perspectives* 9, 185–200.

Milner, A. D. & Goodale, M. A. (1995), *The Visual Brain in Action*, Oxford University Press, New York.

Milner, A. D. & Goodale, M. A. (1998), 'The Visual Brain in Action (Precis)', *Psyche* 4(12).

Mishkin, M. & Ungerleider, L. G. (1982), 'Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys', *Behavioral Brain Research* 6, 57–77.

Nanay, B. (2012), 'Action-oriented Perception', *European Journal of Philosophy* 20(3), 430–446.

Nanay, B. (2011), 'Do we see apples as edible?', *Pacific Philosophical Quarterly* 92(3), 305–322.

Nee, D. E.; Jahn, A. & Brown, J. W. (2013), 'Prefrontal Cortex Organization: Dissociating Effects of Temporal Abstraction, Relational Abstraction, and Integration with fMRI', *Cerebral Cortex (New York, NY)* 24(9), 2377–2387.

Newell, A. & Simon, H. (1976), 'Computer Science as Empirical Inquiry: Symbols and Search', *Communications of the Association for Computing Machinery* 19, 113–126.

Newen, A. & Bartels, A. (2007), 'Animal Minds and the Possession of Concepts', *Philosophical Psychology* 20, 283–308.

Norman, D. A. (1999), 'Affordance, Conventions, and Design', *Interactions*, 38–42.

Ochipa, C.; Rapcsack, S. Z.; Maher, L. M.; Rothi, L. J. G.; Bowers, D. & Heilman, K. M. (1997), 'Selective deficit of praxic imagery in ideomotor apraxia', *Neurology* 49, 474–480.

O'Regan, J. K. & Noë, A. (2001), 'A sensorimotor account of vision and visual consciousness', *Behavioral and Brain Sciences* 24, 939–1031.

Peacocke, C. (1998), 'Implicit conceptions, understanding and rationality', *Philosophical Issues* 9, 43–88.

Pecher, D.; Boot, I. & van Dantzig, S. (2011), Abstract Concepts: Sensory-Motor Grounding, Metaphors, and Beyond, *in* B. Ross, ed., 'The Psychology of Learning and Motivation', Burlington: Academic Press, 217–248.

Perry, J. (1979), 'The Problem of the Essential Indexical', *Noûs* 13(1), 3–21.

Petersen, W. (2007), 'Representation of Concepts as Frames', *The Baltic International Yearbook of Cognition, Logic and Communication* 2, 151–170.

Petersen, W. & Osswald, T. (2012), A Formal Interpretation of Concept Types and Type Shifts, *in* K. Kosecki & J. Badio, ed., 'Cognitive Processes in Language.' Peter Lang, Frankfurt.

Piaget, J. (1977), The Role of Action in the Development of Thinking, *in* W. F. Overton & J. McCarthy Gallagher, ed., 'Knowledge and Development', Springer US, 17–42.

Piaget, J. (1954), *The Construction of Reality in the Child*, Basic, New York.

Piaget, J. (1952), *The Origins of Intelligence in Children*, International Universities Press, New York.

Poincaré, H. (1958), *The Value of Science*, Dover Publications, New York.

Povinelli, D. J. (2000), *Folk Physics for Apes: The Chimpanzee's Theory of How the World Works*, Oxford University Press, New York.

Prinz, J. (2005), The Return of Concept Empiricism, *in* H. Cohenand; C. Leferbvre, ed., 'Categorization and Cognitive Science', Elsevier, Oxford.

Prinz, W. (1997), 'Perception and Action Planning', *European Journal of Cognitive Psychology* 9(2), 129–154.

Proffitt, D. R. & Linkenauger, S. A. (2013), Perception viewed as a phenotypic expression, *in* W. Prinz; M. Beisert & A. Herwig, ed., 'Action science: Foundations of an Emerging Discipline', MIT Press, Cambridge, MA, 171–197.

Pulvermüller, F. (2005), 'Brain mechanisms linking language and action', *Nature Reviews Neuroscience* 6, 576–582.

Pylyshyn, Z. (1973), 'What the Mind's Eye Tells the Mind's Brain: A Critique of Mental Imagery', *Psychological Bulletin* 80, 1–25.

Pylyshyn, Z. W. (1984), *Computation and Cognition*, MIT Press, Cambridge, MA: Bradford Books.

Pylyshyn, Z. W. (1981), 'The Imagery Debate: Analogue Media versus Tacit Knowledge', *Psychological Review* 88(1), 16–45.

Rakoczy, H. (2010), 'From thought to language to thought: Towards a dialectical picture of the development of thinking and speaking', *Grazer Philosophische Studien* 81, 79–105.

Rizzolatti, G.; Camarda, R.; Fogassi, L.; Gentilucci, M.; Luppino, G. & Matelli, M. (1988), 'Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements', *Experimental Brain Research* 71, 491–507.

Rizzolatti, G.; Fadiga, L.; Gallese, V. & Fogassi, L. (1996), 'Pre-motor cortex and the recognition of motor actions', *Cognitive Brain Research* 3, 131–141.

Rizzolatti, G.; Fogassi, L. & Gallese, V. (2000), Cortical Mechanisms Subserving Object Grasping and Action Recognition: A New View on the Cortical Motor Functions, *in* M. S. Gazzaniga, ed., 'The New Cognitive Neurosciences', MIT Press, Cambridge, MA, 539–552.

Rizzolatti, G.; Fogassi, L. & Gallese, V. (2001), 'Neurophysiological mechanisms underlying the understanding and imitation of action', *Nature Reviews Neuroscience* 2, 661–670.

Rizzolatti, G. & Sinigaglia, C. (2010), 'The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations', *Nature Reviews Neuroscience* 11, 264–274.

Rodríguez-Ferreiro, J.; M. M.; Ribacoba, R. & Cuetos, F. (2009), 'Action naming is impaired in Parkinson disease patients', *Neuropsychologia* 47(17), 3271–3274.

Scarantino, A. (2003), 'Affordances Explained', *Philosophy of Science*, 949–961.

Searle, J. R. (1983), *Intentionality*, Cambridge University Press, Cambridge, New York.

Searle, J. R. (1980), 'Minds, Brains and Programs', *Behavioral and Brain Sciences* 3(3), 417–57.

Siegel, S. (2014), Affordances and the Contents of Perception, *in* B. Brogaard., ed., 'Does Perception have Content?', Oxford University Press, New York.

Siegel, S. (2010), *The Contents of Visual Experience*, Oxford University Press, New York.

Simpson, E. A.; Murray, L.; Paukner, A. & Ferrari, P. F. (2014), 'The mirror neuron system as revealed through neonatal imitation: presence from birth, predictive power and evidence of plasticity', *Philosophical Transactions of the Royal Society* 369, 1–12.

Sinigaglia, C. (2010), Mirroring and Understanding Action, *in* Mauricio Suarez; Mauro Dorato & Miklos Redei, ed., 'Epsa Philosophical Issues in the Sciences', Springer, 227–238.

Sirigu, A.; Cohen, L.; Duhamel, J. R.; Pillon, B.; Dubois, B. & Agid, Y. (1995), 'A selective impairment of hand posture for object utilization in apraxia', *Cortex* 31, 41–55.

Smith, B. C. (1986), Varieties of Self-reference, *in* 'Proceedings of the 1986 Conference on Theoretical Aspects of Reasoning About Knowledge', Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 19–43.

Smith, L. B. (2005), 'Cognition as a dynamic system: Principles from embodiment ', *Developmental Review* 25, 278–298.

Sommerville, J. A. & Hammond, A. J. (2007), 'Treating Another's Actions as Ones's Own: Children's Memory of and Learning From Joint Activity', *Developmental Psychology* 43, 1003–1018.

Sommerville, J. A. & Woodward, A. L. (2010), The link between action production and processing in infancy, *in* F. Grammont; D. Legrand & P. Livet, ed., 'Naturalizing Intention in Action', MIT Press, Cambridge, 67–89.

Stoffregen, T. (2003), 'Affordances as Properties of the Animal–Environment System', *Ecological Psychology* 15(2), 115–134.

Stoffregen, T.; Gorday, K. M.; Sheng, Y.-Y. & B., F. S. (1999), 'Perceiving Affordances for Another Person's Actions', *Journal of Experimental Psychology* 25, 120–136.

Strawson, P. F. (1959), *Individuals: An Essay in Descriptive Metaphysics*, Methuen, London.

Synofzik, M.; Vosgerau, G. & Newen, A. (2008b), 'I move, therefore I am: A new theoretical framework to investigate agency and ownership ', *Consciousness and Cognition* 17(2), 411–424.

Synofzik, M.; Vosgerau, G. & Newen, A. (2008a), 'Beyond the comparator model: A multifactorial two-step account of agency ', *Consciousness and Cognition* 17(1), 219–239.

Thelen, E. & Smith, L. B. (1994), *A dynamic systems approach to the development of cognition and action*, Cambridge, Mass. MIT Press.

Tillas, A. (2014), 'How do ideas become general in their signification?', *The Baltic International Yearbook of Cognition, Logic and Communication* 9: Perception and Concepts, 1–35.

Tomasello, M. & Call., J. (1997), *Primate Cognition*, Oxford University Press, Oxford.

Tschentscher, N.; Hauk, O.; Fischer, M. H. & Pulvermüller, F. (2012), 'You can count on the motor cortex: finger counting habits modulate motor cortex activation evoked by numbers.' *Neuroimage* 3139–3148(59).

Tuckman, B. W. & Monetti, D. M. (2010), *Educational Psychology*, Wadsworth, Belmont, CA.

Turvey, M.; Shaw, R.; Reed, E. & Mace, W. (1981), 'Ecological laws of perceiving and acting: In reply to Fodor and Pylyshyn (1981) ', *Cognition* 9(3), 237–304.

Turvey, M. T. (1992), 'Affordances and Prospective Control: An Outline of the Ontology', *Ecological Psychology* 4(3), 173–187.

Ultsch, A. (1993), Self-Organizing Neural Networks for Visualisation and Classification, *in* O. Opitz; B. Lausen & R. Klar, ed., 'Information and Classification', Springer Berlin Heidelberg, 307–313.

Van Gelder, T. (1995), 'What Might Cognition Be, If Not Computation?', *Journal of Philosophy* 92(7), 345–381.

Vosgerau, G. (2009), *Mental Representation and Self-Consciousness. From Basic Self-Representation to Self-Related Cognition*, mentis, Paderborn.

Vosgerau, G.; Seuchter, T. & Petersen, W. (2015), Analyzing concepts in action-frames, *in* Th. Gamerschlag; D. Gerland; R. Osswald & W. Petersen, ed., 'Meaning, Frames, and Conceptual Representation', Düsseldorf University Press, Düsseldorf, 293–310.

Wang, R. F. & Spelke, E. S. (2002), 'Human spatial representation: Insights from animals', *Trends in Cognitive Sciences* 6(9), 376–382.

Weber, A. M. & Vosgerau, G. (2011), Is cognition grounded in action? Conceptual considerations and empirical evidence, *in* B. Kokinov; A. Karmiloff-Smith & N. J. Nersessian, ed., 'European perspectives on cognitive science', New Bulgarian University Press, Sofia.

Weber, A. M. & Vosgerau, G. (2012), 'Grounding Action Representations', *Review of Philosophy and Psychology* 3(1), 53–69.

Weir, A. A. S.; Chappell, J. & Kacelnik, A. (2002), 'Shaping of Hooks in New Caledonian Crows', *Science* 297(5583), 981.

Weir, A. & Kacelnik, A. (2006), 'A New Caledonian crow (Corvus moneduloides) creatively re-designs tools by bending or unbending aluminium strips', *Animal Cognition* 9(4), 317–334.

Willatts, P. (1985), 'Adjustment of means-ends coordination and the representation of spatial relations in the production of search errors by infants', *British Journal of Developmental Psychology* 3(3), 259–272.

Wilson, M. (2002), 'Six views of embodied cognition', *Psychonomic Bulletin & Review* 9(4), 625–636.

Wimmer, H. & Perner, J. (1983), 'Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception', *Cognition* 13(1), 103–128.

Wimmer, H.; Wachter, J. & Perner, J. (1982), 'Cognitive Autonomy of the Development of Moral Evaluation of Achievement', *Child Development* 53(3), 668–676.

Wittlinger, M.; Wehner, R. & Wolf, H. (2006), 'The ant Odometer: Stepping on Stilts and Stumps', *Science* 312, 1965–1967.

Xu, F. & Carey, S. (1996), 'Infants' Metaphysics: The case of numerical identity', *Cognitive Psychology* 30, 111–153.

Zwaan, Rolf A.; Zwaan, R. A. (1999), 'Embodied cognition, perceptual symbols, and situation models', *Discourse Processes* 28(1), 81–88.