Discussion
Towards a common framework of grounded action cognition: Relating motor control, perception and cognition

Antje Gentsch a,b, Arne Weber c, Matthís Synofzik d,e, Gottfried Vosgerau c, Simone Schütz-Bosbach b,f,*

a University College London, CEHP Research Department, London, UK
b Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
c Department of Philosophy, Heinrich-Heine-University Düsseldorf, Germany
d Department of Neurodegenerative Diseases, Hertie-Institute for Clinical Brain Research, University of Tübingen, Germany
e German Research Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany
f Psychology Department, Ludwig-Maximilians-University, Munich, Germany
*Corresponding author at: Psychology Department, General and Experimental Psychology, Ludwig-Maximilians-University Munich, 80802 Munich, Germany.
E-mail address: simone.bosbach@gmail.com (S. Schütz-Bosbach).

A R T I C L E   I N F O
Article history:
Received 8 May 2015
Revised 11 September 2015
Accepted 14 September 2015
Available online 24 September 2015

Keywords:
Grounded cognition
Motor control
Action perception
Action cognition
Simulation
Internal models
Common coding

A B S T R A C T
The relation between motor control and action cognition – including action-related thoughts and action-related perception – has been subject to controversial discussions in the last three decades. During these decades, cognitive neuroscience has been increasingly confronted with a huge variety of different accounts trying to understand and explain the relation between these systems, their interdependencies and the mediating mechanisms by establishing notions such as “internal models”, “simulation” or “shared representation”. These accounts, however, include a large array of partly overlapping, partly contradictory theories using similar terms for different mechanisms and different terms for similar mechanisms. In the absence of a systematic work-up and comparison, this array of accounts and theories leads to confusion in the field, duplication of experimental work, and unconnected parallelism of theory formation within and between different disciplines. Here we provide a systematic comparison of current models and prospective theories that deal with the relation between cognition, perception and motor control mechanisms. In a second step, we propose “grounded action cognition” as a comprehensive metatheoretical framework which defines different hypothetical possibilities of the relations between these domains, offers systematic insights into current models and theories and last but not least may help to increase comparability of empirical research in the domain of action and action cognition.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

There is a long tradition and interest in many different disciplines on the interplay between action and perception. This interest is led by various aims: while some disciplines investigate this interplay with the aim of understanding human ‘cognitive mechanisms’ (cognitive psychology, philosophy), brain mechanisms (cognitive neuroscience) or neurologic and psychiatric diseases (neurology/psychiatry), others study this relation to model skillful movements (robotics) or perceptual abilities (computational and mathematical neuroscience). Some of these traditions date back to 19th century experimental psychology and philosophical theory of voluntary action (James, 1890/1981). Others were developed more recently. For example, approaches from computational and cognitive neuroscience, starting in the early 1990s, aim at describing the precise brain systems and neuronal dynamics underlying the action–perception inter-linkage by means of internal forward/generative models (Friston, Daunizeau, Kilner, & Kiebel, 2010; Wolpert & Miall, 1996). The focus on the linkage between action and perception has now been expanded to include also conceptual abilities and cognition in general. Specifically, growing interdisciplinary work has now begun to relate different theoretical approaches and empirical findings to explain also higher-cognitive skills like mind reading in social contexts (Gallese & Goldman, 1998) or mental imagery (Grush, 2004).

Here we take an interdisciplinary point of departure with the aim to provide a systematic comparison of current and established theoretical models and prospective theories that deal with the relation between cognition, perception and motor control mechanisms. We will in particular focus on the proposed internal representational mechanisms governing mutual relations between perception and action. Accordingly, the selection criterion for the theories to be compared in the present paper is that they make
some substantial claim about the systematic connection between the domains of cognition, perception and motor control.\(^1\) It is beyond the scope of the present endeavor to also take dynamic interactions between mind/brain, body and world into account as proposed by some more radical conceptions of embodied cognition (for an overview compare Shapiro, 2011; or Wilson, 2002; for certain conceptions of these radical or dynamic views see Beer, 1995; Hutten & Myin, 2013; Keijzer, 2002). Thus, we will here specifically consider internal representations that appear to draw on motor-related processes. We will use the term ‘grounding’ to refer to the general relation between motor processes and action perception/cognition, which seems underdetermined so far in current theories. Please also note that it is not our aim to provide a systematic review of those single theories in terms of supporting or challenging empirical evidence (for a recent review see, for example, Engel, Maye, Kurthen, & König, 2013). Instead, we here aim to systematically compare the empirical and explanatory foci those theories adopt with respect to the nature of the relation between motor control and action-cognition.

This paper comprises three sections. The first provides a brief overview of different action theories and suggests a classification by target mechanisms. In the second section, we introduce a framework that will be used to illustrate our notion of grounded action cognition as a metatheoretical view. Third and finally, we try to show how existing theories can or cannot be classified into genuine grounding theories. This new classification scheme shall offer new perspectives into commonalities and differences as well as the explanatory scope with respect to the degree to which perceptual and cognitive abilities genuinely draw on motor capacities; however, it is not meant to suggest an evaluation of existing and established theories concerning their theoretical and empirical adequacy.

2. The three main families of action theories

The large amount of highly heterogeneous, partly overlapping, partly differing theories coming from very different disciplines is often confusing. To facilitate a better overview and understanding, we here classify the existing theories into three major theoretical frameworks/families of grounded action cognition accounts: (1) Common Coding, (2) Internal Models, and (3) Simulation theories (Table 1; for a detailed description and analysis of each of these families, see next paragraph). These theories operate at different levels and interfaces between domains of action cognition and use diverse conceptual tools (see Table 1). One shared assumption of all major theories is that some kind of common representational ground for action and cognition exists. This assumption has been developed, however, at different levels of action cognition, and with crucial differences in the empirical and explanatory focus on motor control, perception or conceptual abilities (see Fig. 1). This includes fundamental differences in the abilities these theories try to account for: while some theories resort to a common ground between action and cognition to explain action planning or motor imagery, others refer to this notion to explain the understanding of other individuals’ actions. The main assumptions regarding the central thesis of grounded action cognition are summarized in Table 1.

In order to understand the differences between the theories in their attempt to link the domains of action, perception and cognition, it is helpful to consider the primary domain a theory was originally developed for and how it has been extended. For example, common coding theories as well as some internal model accounts, such as motor control theory, were explicitly developed from a motor control point of view. Therefore, the theoretical concepts are grounded in notions of efferent and (re-)afferent information (i.e., motor output and sensory feedback) which have been applied not only to understand the mechanisms behind movement kinematics (motor control theory) but also to describe the semantics of action (ideomotor theory). In contrast, another version of internal models, predictive coding, was originally developed to explain visual perception and its underlying computational architecture. It has been extended to other sensory modalities as well as to action in terms of active inference (the latter being a more recent development). Hence, the key concepts in this model are rooted in assumptions around the organization of sensory input and the perception of the outside world. The ontogeny of simulation theories, the third main family of action theories, is most heterogeneous, since the notion of ‘simulation’ has been used to account for a variety of phenomena in the domains of action cognition and perception. For example, the notion of simulation has been used to explain a perceptual understanding of other’s actions (mirror neuron theory) and the ascription of mental states to others (simulation theory), as well as the imagination of action (motor imagery) and even abstract conceptual abilities (perceptual symbolic theory). Hence, understanding the primary phenomenon that each theory is trying to account for is crucial for appreciating how existing theories differ in relating the domains of cognition, perception and motor control mechanisms. The classification of the three families of theories according to their primary domain of explanation is illustrated by Fig. 1.

3. Grounded action cognition: a metatheoretical view

In our approach towards a systematic framework, we will now specify common denominators of the manifold existing assumptions regarding the relation between the three domains of action cognition, perception, control and concepts of action. Current theories and related experimental reports sometimes appear to use the same terms to explain different phenomena or mechanisms, and different terms to explain the same phenomena or similar mechanisms. For example, the term ‘simulation’ has been used to refer to sensorimotor processes as a reactivation of neural motor circuitries not only when observing actions but also in the context of reasoning about mental states of others. Likewise, the finding of an activation of cortical areas involved in motor control during action observation has been interpreted and labeled interchangeably as ‘mirroring’, ‘simulating’ or ‘mentalizing’. Moreover, terms such as ‘embodiment’, for instance, have been applied in a narrow sense with reference to self-awareness as a form of bodily awareness (i.e., the embodied self) but also in a broader sense to explain a fully functional system of abstract concepts (cf. below, perceptual symbol system theory).

Moreover, study reports sometimes use vague or even overstated wording when reporting their findings and interpreting the relation between the three domains. For instance, authors sympathetic with simulation theories often use various and quite strong expressions like cognition/perception (x) is “based” on a sensorimotor process/the sensorimotor system (y), x “arises” in y, x is “localized” in y, x is “derived” from y, x is “specified” by y. x is

\(^1\) For this reason, the sensorimotor theory of visual consciousness by O’Regan and Noë (2001), for example, is not considered in the present framework: Although it claims that visual perception is a form of action, it does not propose a specific linking mechanism between perception and action (unlike ideomotor theory, for example).
Systematic overview of theories of action cognition. These theories assume a tight, but differing interrelation between the domains of motor control, action perception and action cognition. They can be grouped into three different families of grounded action cognition accounts: Common Coding, Internal Models, and Simulation theories.

<table>
<thead>
<tr>
<th>Theory</th>
<th>Explanatory target</th>
<th>How do motor cognition, motor control and action perception relate?</th>
<th>Origin</th>
<th>Key references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event Coding</td>
<td></td>
<td>A common code: the representational format for distal perceptual contents and motor plans is identical and formed through principles of temporal coupling and synchronized activation</td>
<td>Cognitive Psychology (action perception)</td>
<td>Hommel et al. (2001) and Prinz (1997)</td>
</tr>
<tr>
<td>Internal Models</td>
<td></td>
<td>An internal forward model which represents the dynamics of the motor system, including all sensory states resulting from motor actions, in order to predict future states</td>
<td>Computational Modeling, Cognitive Robotics (motor control)</td>
<td>Wolpert and Miall (1996)</td>
</tr>
<tr>
<td>Motor Control Theory</td>
<td>Implementation of optimal online control of movements and posture</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predictive Coding</td>
<td>Sensory perception &amp; motor control Perception as prediction</td>
<td>Active inference under the free-energy principle: sensory surprise automatically translates into motor control representations (similar to motor reflexes) that serve to minimize sensory prediction errors by guiding sensory sampling</td>
<td>Computational Modeling (perception in general)</td>
<td>Friston (2005) and Friston, Mattout, and Kilner (2011)</td>
</tr>
<tr>
<td>Emulation Theory</td>
<td>Motor control &amp; motor imagery</td>
<td>Draws on and combines elements of internal models and mirror mechanisms central to motor control theory and mirror neuron theory</td>
<td>Philosophy (action cognition)</td>
<td>Grush (2004) and Hurley (2008)</td>
</tr>
<tr>
<td>Simulation Theories</td>
<td>Mind reading</td>
<td>The concept of mental simulation: action observation (i.e., the perceptual representation of a motor act of others) involves the activation of mental state sequences that are isomorphic to those underlying self-generated actions</td>
<td>Philosophy, Cognitive Neuroscience (social cognition)</td>
<td>Gallese and Goldman (1998), Goldman (1992), and Gordon (1986)</td>
</tr>
<tr>
<td>Mirror Neurons</td>
<td>Neuroanatomy of observed actions of others</td>
<td>The notion of mirror neurons: a population of neurons coding for both observed and executed actions, that is, the same neural system is used for representing motor acts and motor percepts</td>
<td>Neurophysiology (action perception)</td>
<td>Gallese et al. (1996) and Rizzolatti, Fadiga, Gallese, and Fogassi (1996)</td>
</tr>
<tr>
<td>Perceptual Symbols</td>
<td>Conceptual thinking</td>
<td>The construct of simulators: Concepts act as simulants by relying on a re-activation of the sensorimotor states involved in experiencing the actual entity</td>
<td>Cognitive Psychology (cognition in general)</td>
<td>Barsalou (1999) and Barsalou, Kyle Simmons, Barbeay, and Wilson (2003)</td>
</tr>
<tr>
<td>Theory</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

"embodied"/"grounded" in y (passim). x is “constituted” by y. x “resides” in y. x/y “integrate(s)” y/x. x/y “enters” into x, y is “engaged” in x, or y “essentially is” for x. All of these formulations are used to interpret data from experimental work usually showing a correlation between two factors. However, the fact that two factors correlate can be interpreted in very different ways, from mere unspecified modulation of one factor by the other to causal dependence of the one factor on the other. Here we propose grounded action cognition as a metatheoretical framework that aims to define hypothetical possibilities for the relations between different domains of action representation in order to provide a systematic synopsis and a path towards an integrative understanding of current theories of action. Within this new framework, a crucial distinction is drawn between constitution relation and acquisition relation, which offers a systematic classification tool for existing and future theories of grounded action cognition.

3.1. Acquisition and constitution: specifying the relation between perception/cognition and action

In order to specify the claim of grounded cognition, it is crucial to scrutinize the exact relation that is assumed to exist between two domains of interest (e.g., between motor control and action perception, or action cognition). So far, the question of whether grounded or embodied cognition is meant to oppose “classical” modular pictures of the mind (Fodor, 1975; Pylyshyn, 1984; Smith & Medin, 1981; Tulving, 1972) is in fact not answered. There is certainly no doubt that action (i.e., motor control) and perception/cognition are somehow interrelated. The question is thus not whether these processes are related, but how they are related. Representatives of the “classical” picture of the mind even do not...
However, they also explicitly deny a perception–action interrelation (although they usually do not focus on it), but they typically claim this relation to be rather weak, for example in the sense of mere modulating bi-directional influences. Thus, our reconstruction of the perception–action interrelation meant by the different authors above will be such that their claims are indeed opposing more classical views in claiming a much stronger, thus functional determining relation between action and perception in the sense of acquisition and/or constitution. In a first step, we will now introduce the technical terms of acquisition conditions and of constitution, respectively (cf., Weber & Vosgerau, 2012). We will illustrate how we understand these terms with the help of empirical examples and we will also discuss some empirical implications that follow from this new perspective.

Acquisition can be illustrated by a classical experiment of Held and Hein (1963) in newborn animals. They divided kittens into two groups; kittens in one group pull a carriage around a room in a horizontally rotating carousel, while kittens in the other group ride this carriage. The carousel was set up such that both groups of kittens have the same visual stimuli, the only difference being that kittens in the first group are allowed to make active movements while the kittens in the other group were only passively moved. After spending some time in the carousel, tests show that the riding kittens did not develop normal perceptual abilities as for example depth perception. Obviously, these kittens were not able to integrate their visual experiences with movements of their own bodies, which caused visual deficits in the long run (see also, Gibson, 1968). These findings suggest that the ability to actively move (ability A) is among the acquisition conditions for the ability to perceive depth (ability B). However, once B is acquired on the basis of A, ability A can get lost without any negative impact on ability B. So, a kitten that already acquired normal sight will not lose its ability to perceive if we put it in a carriage (thus hindering it from active movement). Active movement is only necessary to acquire the ability to perceive but not necessary to maintain this capacity once it is fully acquired.

"Constitution", in contrast, refers to a much tighter interpretation of the relation between action and perception. An almost trivial example to illustrate a real constitution relation is that the ability to move one's own legs is constitutive for the ability to walk, while both abilities are not identical, since it is possible to be able to move one's own legs without being able to walk. So, already in this example it is obvious that constitution does not necessarily mean that there is no more to ability B than just ability A: indeed, there is much more to the ability to walk than just the ability to move one's own legs. Nevertheless, you cannot do the latter without being able to do the former. However, constitutive relations can also be interpreted less strong. For example, if the complete loss of A only leads to some impairments of B, it is still valid to argue that A is a constitutive element of full-blown ability B. Thus, although the criterion for constitutive relations might read very strong at first sight, it only requires a systematic dependence between break-downs in either ability.

So far only relatively few empirical studies were able to indicate that motor control processes indeed play a constitutive role in
action cognition, thus going beyond mere correlations between motor control processes and action cognition. A prominent example is a study that investigated patients diagnosed with amyotrophic lateral sclerosis (ALS). ALS refers to a neurodegenerative condition that is defined by progressive degeneration of motor tracts. This study showed that ALS patients were significantly more impaired in concepts of actions than in concepts of objects, and that this difficulty specifically correlated with atrophy of the motor cortex (Grossman et al., 2008). These findings suggest that intact conceptual knowledge of action features depends on an intact motor cortex. If this finding can be corroborated by future studies, confirming indeed a specific relation between degraded action knowledge and degraded motor systems (and not just degraded action knowledge as part of the more widespread extramotor cortical atrophy known in ALS (Libon et al., 2012), this would suggest that motor systems are indeed constitutive for action cognition. Intact motor systems would here present the structural presupposition to intact action cognition.

In another domain, intact motor processes present the functional presupposition for acquiring intact perception. Microsaccades, for example, allow restoring faded vision during fixation, for both foveal and peripheral targets. Accordingly, it has been shown already in the 1950s that when eye movements are completely suppressed (for instance, by techniques of retinal stabilization (Ditchburn & Ginsborg, 1952), our perception of stationary objects appears to fade completely (Martinez-Conde, Macknik, Troncoso, & Dyar, 2006; for review see, Martinez-Conde, Otero-Millan, & Macknik, 2013). For the more specific thesis that motor processes are indeed constitutive elements for acquiring intact action perception, it would be necessary to demonstrate that similar process are at work in the perception of actions.

3.2. The partial constitution criterion

Our proposal here is that the only genuine grounding relation is that of a partial constitution relation between perception and action, in the sense that there is more to ability B than just ability A. Thus, on this view the relation does not collapse into an identity relation. To make this argument clearer, let us take the hypothetical case of a radical and full constitution of a certain perceptual/cognitive ability by motor processes. This would lead us to the assumption that this ability is just a subclass of or even identical to some (complex) motor ability. Moreover, if all perceptual/cognitive abilities would be so fully constituted by motor processes, some basic motor abilities would be constitutive of all action-perceptual and -cognitive capacities, and this implies that the two domains would fall together and become the same. However, if perceptual processes or states are just kinds of motor processes or states, there would be no intelligible sense of understanding a certain perceptual process as “grounded” in another basic process, since then there is nothing left to be “grounded”; the former process would be merely identical to the latter. Moreover, the different domains of motor control and action perception would not be discriminable in empirical research. The other hypothetical extreme case would be that there is no constitution at all of one process by another. This would then in the end lead to a picture similar to the classical conception of a modular mind (see above). Then, however, there would be no novel and challenging way of talking about some cognitive processes potentially being grounded. Such relations would simply be too weak to establish a picture opposing the classical one. Therefore, we argue that grounded action cognition should be understood and specified in terms of a partial constitution relation. The different types of constitution relation are illustrated by Fig. 2.

4. Action cognition theories and implied constitution relation

In the first part of the present paper, we organized the established action cognition theories into three major families, Common Coding (CC), Internal Models (IM) and Simulation Theories (ST). We will now analyze them in more detail with respect to whether they can be considered genuine grounding theories; that is, they will be compared according to the implied constitution relation between motor control and action cognition. The aim shall be to put the variety of existing accounts of action into a systematic frame of grounded action cognition. Each family contains some theories that implicitly assume an at least partial constitutive relation between action and perception, while other theories in contrast do not imply a constitutive relation. In Table 2, we made an attempt to organize those theories according to the constitution relation they assume to hold. We will discuss the categorization of each theory family in our table in turn.

4.1. Common Coding Theory

Common Coding theories are characterized by the assumption that there is a bidirectional or common sensorimotor code. The ideomotor theory assumes that action and perceptual representations are separate representations in principle, but can be co-activated to generate and facilitate voluntary action and adaptive control of goal-directed behavior. Thus, the motor domain is here not seen as constitutional for perceptual and cognitive abilities (see Table 2). In contrast, the theory of event coding holds that there is a common representational code for action and perception. This code structures action planning and action perception and therefore action and perception are related through this

![Fig. 2](image-url). Schematic representation of different types of constitution relation between two abilities, A and B. (i): Some B-ability correlates with some A-ability (dotted line), while there is no constitution relation between the two. (ii): A B-ability is partially constituted by A-abilities if the loss of A leads to severe impairment of the B-ability but not to a complete breakdown (since there is more to ability B than just A). (iii): The B-ability is fully constituted by A-abilities if there is no non-A-ability constitutive for it, and in this case the B-ability is nothing but a (complex) A-ability itself.
The three major families of action cognition theories and the main assumptions regarding conditions of constitution.

<table>
<thead>
<tr>
<th>Constitution relation</th>
<th>Theory</th>
<th>Central hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Coding</td>
<td>No</td>
<td>Ideomotor theory</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>Theory of event coding</td>
</tr>
<tr>
<td>Internal Models</td>
<td>Partial</td>
<td>Motor control theory</td>
</tr>
<tr>
<td></td>
<td>Partial</td>
<td>Predictive coding</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>Emulation theory</td>
</tr>
<tr>
<td>Simulation Theories</td>
<td>Full</td>
<td>Mirror neuron theory; Motor imagery</td>
</tr>
<tr>
<td></td>
<td>Full</td>
<td>Perceptual symbols theory</td>
</tr>
</tbody>
</table>

representational format. Critically, however, the code is independent of executive action domains as well as of sensory modalities. Thus, following this picture, the domain of motor control is not directly connected to (or even constitutive for) sensory modalities. Instead, a common code enables an indirect communication between the two domains. Moreover, motor control is not seen as the primary system because the perceptual system and the motor control system are coequal.

The main aim of Common Coding theories is to explain the representations of action goal and action effects (a common representational domain for perception and action) with the cognitive antecedent of the action (prospective coding in event representation; Schütz-Bosbach & Prinz, 2007). The main thesis is that features relevant for motor control are involved already at the level of action perception and the perception of action effects, with abundant empirical evidence, for example from classical studies on spatial stimulus–response compatibility effects (e.g., Elsner & Hommel, 2001; Hommel, Musseler, Aschersleben, & Prinz, 2001; Kunde, 2001; see Shin, Proctor, & Capaldi, 2010 for a review). The typical observation is an advantage (in terms of speed and accuracy) of responses that share “motoric” features with a target stimulus (such as the relative location in space), supposedly due to the automatic activation of a common location code between stimulus and response. An elaborated version of the theory of event coding assumes a common format of (late) perception and (early) motor control which are coded in a common representational medium, i.e. event codes that are both motoric and perceptual, commonly coded, “especially if the features of perceived and to-be-produced events overlap […] In other words, perceptual and action-planning processes only interact if the codes they operate on refer to the same (kind of) feature of a distal event” (p. 862, Hommel et al., 2001).

In contrast to Simulation Theories (as in the case of mirror neurons), Common Coding – at least in its original version – is not a framework for understanding actions of others but mainly for action planning. In contrast to Internal Model theories, action control in Common Coding does not draw on a forward signal which is sent in parallel with the motor command to guarantee optimal action performance. Whereas Common Coding emphasizes perception because it supposedly shapes motor control (at least with respect to distal events) Internal Model theories focus primarily on motor control. To summarize, Common Coding assumes three different classes of representational formats: purely motoric, purely sensory and a common code format. Thus, neither motor control is grounded in perception nor the other way round, but both are linked through additional (common code) representations. None of the Common Coding theories can therefore be considered, from a constitution point of view, as a genuine grounding theory.

4.2. Internal Models

Internal Model theories are characterized by the proposed existence of internal models of the motor apparatus. These models are postulated neural processes that extrapolate the commands of the motor system in order to estimate and anticipate the outcome of a motor command. One variant of Internal Model theories, motor control theory, has its origin in control theory and robotics and primarily aims at explaining optimal online control of movements such as posture control and movement trajectory planning. More precisely, motor control theory assumes different internal models: a forward internal model predicts the sensory consequences from efference copies of issued motor commands; additionally, an inverse model calculates the necessary feedforward motor commands from the desired movement trajectory (cf. Kawato, 1999). To explain perception, especially of self-generated actions, forward models are sometimes also simply referred to as “effference copy” (von Holst & Mittelstaedt, 1950) or “corollary discharge” (Bell & Grant, 1989; Sperry, 1950). Following this interpretation, perception and action are essentially dependent on motor intentions and related motor signals, such as the effference copy and predictions computed on their basis. Extensive evidence for this dependency comes from studies on the phenomenon that people cannot tickle themselves, for which predictions have been attributed a prominent role by attenuating the sensory consequences of self-generated actions (Blakemore, Frith, & Wolpert, 1999). However, the internal model itself, despite being a decisive part of motor control, is not identical with the motor mechanism, otherwise a comparison between a predicted and actual state for optimizing motor control (e.g. fast online corrections of movement trajectories) would not be possible. That is, motor control theory postulates that there is one input signal (i.e. desired action outcome) and two different output processes (i.e. motor command and effference copy), and motor control is emphasized as a constitutive condition. Therefore, according to our suggested framework, a constitution of perception by motor control mechanisms can be found but this constitution is not exhaustive (i.e. there are also other factors involved). Hence, in motor control theory, some motor representations are considered partially constitutive at least for parts of action cognition.

Another type of Internal Model theory, predictive coding, is based on the view of perception as prediction (Friston et al., 2010). In contrast to motor control theory, the primary focus of this approach is on perceptual models and not on motor control mechanisms. However, active inference as a corollary of this perspective also aims to account for many aspects of motor behavior, from oculomotor reflexes to goal-directed movements. It assumes a generative forward model which maps from motor control to sensory consequences, and which forms part of a larger hierarchical model.
used for perceptual inference in general. In contrast to motor control theory, it is thought to serve all aspects of motor control without the need for a separate inverse model to provide control signals (cf. Friston, 2005). In this respect, forward-inverse models in motor control are formally distinct from predictive coding accounts of action perception. In predictive coding accounts, motor commands are not constructed by an inverse model but arise from an inversion of the forward model. That is, instead of a mapping from control signals to sensory consequences and vice versa, motor control signals are identical with descendancy sensory predictions of proprioceptive sensations. However, beyond internal proprioceptive consequences of movements, exteroceptive consequences of actions (e.g., vision, audition) are considered corollary discharges, as in classical forward models. Thus, at least for exteroceptive modalities of action perception in predictive coding mechanisms, the same holds true as for motor control theories: some motor representations are constitutive for parts of action cognition.

Some types of Internal Model theories incorporate elements from Simulation theories. For example, the emulation theory, as proposed e.g. by Rick Grush (2004), combines constructs like inference copy and motor prediction with mirror mechanisms. It seeks to explain how motor control operations in classical motor centers can produce motor imagery and certain aspects of visual perception by driving the operation of a motor-visual emulator of the musculoskeletal system. Beyond the motor domain, the theory holds that imagery in general is based on both amodal and modality-specific systems driving parallel emulators as real time models of behavior and perception. Hence, the theory has been applied not only as an explanatory model of motor control and motor imagery, but also of several other cognitive functions, indicating dissimilarity between motor processes simpliciter and cognitive functions. Accordingly, in Grush’s discussion of emulation versus simulation accounts of motor imagery, he insists that “mere operations of the motor center is not enough: to produce imagery the motor centers must be driving an emulator of the body” and “[a] motor plan is one thing, a sequence of proprioception and kinesthesia is another” (p. 385; Grush, 2004). Thus, according to Grush, the executed action and its effects are not sufficient (and maybe not even necessary) to install action-related perception and cognition.

Taken together, all types of Internal Model theories postulate that action perception and cognition are significantly informed and construed by predictions. In particular, exteroceptive predictions are not identical with motor commands but are grounded in the latter. However, more than motor commands is necessary to produce these predictions, namely sensory feedback, context cues, and context estimates. Thus, predictions and related action perception/cognition cannot be exhaustedly grounded in motor commands. Therefore, according to Internal Model theories, motor abilities would be (at most) partially constitutive of action-related perceptual and cognitive capacities.

4.3. Simulation Theories

The family of Simulation Theories is characterized by the general assumption that cognition is essentially simulation-like (re-)use of the brain’s modal systems, such as the sensorimotor system. For example, the mirror neuron theory assumes a population of neurons coding for both observed and executed actions, that is, the same neural sub-system is used for representing motor acts and motor percepts (for review see, Rizzolatti & Craighero, 2004). This applies specifically to so-called ‘strictly congruent’ mirror neurons (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996), encoding the kinematics of an action, in contrast to some class of ‘broadly congruent’ mirror neurons (or ‘logically-related’ mirror neurons; e.g. Iacoboni et al., 2005) which encode action goals with a less strict correspondence between motor control and action perception. It should be noted that mirror neurons are regarded as neurophysiological instantiations of simulation and, therefore, mirror neuron theories provide a mechanistic explanation of the involved neuronal processes, rather than offering a theoretical account on their own.

Similarly, the motor imagery account states that motor perceptions (or “motor images”) are “covert actions”, i.e. a simulation of neural mechanisms and dynamics of real, physical action in low level motor centers before an action is executed (Jeannerod, 2001). Here, Jeannerod (2001, p. 106) argues that “lesions affecting structures located ahead of motor system […] affect the ability to produce a given S-state”. By the term S-state, Jeannerod refers to “those ‘mental states’ which involve an action content and where brain activity can be shown to simulate that observed during the same, executed action” (Jeannerod, 2001, p. 103). But these simulations are “covert actions” that do not result in muscular activity as an overt movement.

Another prominent example of Simulation Theories is the theory of perceptual symbol systems (Barsalou, 1999). According to this theory, cognition exclusively consists of activation patterns of various sensory modalities, here the perceptual symbol system. In this case cognitive and perceptual mechanisms share the same representational states because cognitive processing is essentially a simulation, i.e. a re-activation of sensory states, including sensorimotor activation patterns underlying the original perceptual state. This reactivation of the same neural circuits involved during perception of and action towards those objects allows the reenactment of multimodal information. Consequently, also higher-order abilities such as thinking or categorization are “grounded” in low-level sensorimotor abilities. Empirical evidence for the theory of perceptual symbol systems indicates that perceptual and motor mechanisms are engaged when a cognitive system performs conceptual processing. It is suggested that, for example, talking or thinking about objects actually (partly) consists of the reactivation of previous experiences stored in the perceptual symbol system (for an overview cf. Barsalou, 2008; de Vega, Glenberg, & Graesser, 2008; Pecher & Zwaan, 2005; Semin & Smith, 2008).

Based on mirror neuron theories, even cognitive abilities such as mind-reading and understanding mental states of other’s by observing their actions are often “explained” by simulation mechanisms that shall help “put oneself in the ‘mental shoes’ of others” (cf. Gallese & Goldman, 1998).

All variants of simulation theories understand sensorimotor processes as being primary (in contrast to Common Coding). Moreover, in contrast to Internal Models, they focus on higher order cognitive abilities to explain conceptual thinking and mind-reading underlying social understanding and social communication, rather than focusing solely on motor control functions. Following Simulation Theories in general, cognitive abilities such as motor imagery, conceptual abilities or understanding the mental states of others rely on a reactivation of sensorimotor states. Because simulation is essentially a reenactment of sensorimotor modalities there is nothing over and above those sensorimotor mechanisms. Thus, simulation theories postulate an identity relation (i.e., full constitution) between motor abilities and action cognition. A meaningful use of the notion of ‘grounded cognition’, however, can only be based on a relation of constitution and not of an identity between the relevant domains of action cognition, perception and motor control, as claimed by simulation theories.

5. Conclusions

Following the modular picture of the mind, cognition and action has commonly been assumed to operate in strictly different
domains: cognition draws on abstract conceptual representations, whereas motor control functions are thought to rely on relatively low-level and automatic processes. In recent years, however, cognitive neuroscience has been confronted with an increasingly large array of partly overlapping, partly contradicting theories that diverge broadly with respect to the degree to which cognition is thought to be grounded in systems for perception and action. Consequently, the proposed conceptual relations of motor control and action cognition are manifold. This has led to apparent duplication of experimental work, a relative huge collection of unsorted experimental findings and unconnected development of different theories. Recently, there have been some attempts comparing and linking different theories pairwise. For example, connections have been made between theories belonging to the family of Internal Models and Simulation Theory (Iacoboni, 2009), or between Common Coding and Internal Models (Chambon & Haggard, 2013). However, up to date, a comprehensive overall framework able to accommodate all the theories is still missing. We here argued that “grounded action cognition” investigating constitution conditions between motor processes and action cognition may provide a comprehensive meta theoretical framework that allows bringing together this heterogeneous mix of theories and to analyze and classify them according to the specific relation they claim to hold between motor processes and action perception/action cognition. Thereby, we do not evaluate the different views according to their theoretical and empirical adequacy as such, but we do evaluate them according to their claim that there is a grounding relation between motor processes and action perception/cognition that is denied by “classical” theories. Thus, some of the theories turn out to be genuine grounding theories, while others do not qualify as grounding theories. Theories that do not qualify as grounding theories despite of their wording do so either because they claim a grounding relation that is too weak to contradict classical theories, or because they essentially claim that the domains are identical, and thus overstretch the relation for grounding.

Of the nine different action theories, which we classified into three different families based on their primary explanatory scope, only very few can be considered genuine grounding theories. According to Common Coding accounts, separate representational codes are assumed which mediate between action and perception and thereby subserve both the control and evaluation of actions. However, acting and perceiving are seen as independent from each other, hence they are without any constitutive function for the other domain. Internal Model theories take the grounding relation one step further by assuming that perceived action events are essentially construed by predictions. These predictions are thought to be grounded in motor commands, but they are also informed by other types of contextual cues and prior knowledge. Thus, implicit in the idea of internal models is the assumption that motor abilities are partially constitutive of action-related perceptual and cognitive capacities. Within the family of Simulation Theory accounts, the classification depends on the specific version. For example, the mirror-neuron version supposes an identity relation between motor control and perceptual processes in the early stages of processing. Similarly to Internal Model theories, the emphasis is on motor control mechanisms, however, these are understood as fully constitutive for some perceptual processes. In more general versions of Simulation Theory – like Barsalou’s perceptual symbol system – the sensorimotor system plays a constitutive role for all cognitive and perceptual processing, the constitution being exhaustive.

In summary, considering action theories from a grounded cognition perspective provides a systematic tool for evaluating theoretical assumptions on how perception and action are represented and shape one another. This may help to determine adequate empirical approaches to core questions around embodied human cognition. For example, to which extent can motor control deficits or motor constraints affect the perception of other’s actions and socio-cognitive functions at higher levels? How does motor expertise or training determine action-related perceptual content? Which aspects of self-awareness in the domain of action, such as feelings and judgments of self-agency, are reducible to certain motor control functions? The present framework of constitution and constitution relations opens a new perspective for specifying the relation between motor processes and action cognition and may help to find the right balance between reductionist point of views on the one hand and purely associative theories on the other hand. It may help to further discriminate the different domains of action perception and motor states as well as gain further insights in common factors and interdependences of different domains.

Acknowledgments

This project was supported by a grant of the Volkswagen Stiftung to all of the authors (“Grounding thoughts in actions, Thin-kAct”; Az. II/85 068/85 158/85 159) and a Heisenberg fellowship (SCHU 2471/4-1) of German Research Foundation/Deutsche Forschungsgemeinschaft (DFG) to SS-B.

References