



UvA-DARE (Digital Academic Repository)

Sculpting the space of actions: explaining human action by integrating intentions and mechanisms

Keestra, M.

Publication date
2014

[Link to publication](#)

Citation for published version (APA):

Keestra, M. (2014). *Sculpting the space of actions: explaining human action by integrating intentions and mechanisms*. [Thesis, fully internal, Universiteit van Amsterdam]. Institute for Logic, Language and Computation.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

4 THE BRAIN AS A MECHANISM CAPABLE OF KLUDGE FORMATION AND OPEN TO EXTERNAL INFORMATION

We began this part by noting that development and learning allow a baby to expand its vocal repertory from mere babbling without clear rhythm and distinct melody to a wide range of vocal expressions. At the other end of the spectrum, an expert singer can learn complete opera roles that require extensive singing and acting performance simultaneously. According to the neuroconstructivist account of development and learning, the mastery of such vocal skills corresponds with a process of modularization. This process is associated with both the proceduralization or automatization of these skills and the explicitation or increasing capability to articulate the relevant representations, allowing an expert to adjust and correct his cognition and behavior (Karmiloff-Smith 1992 ; Mareschal, Johnson et al. 2007).²²⁴ Related to these processes are – usually implicit - processes that result in shifts between types of information processing, aiming to mitigate the consequences of capacity limitations in the brain, like the strategies of segmentation or chunking of information (Halford, Wilson et al. 1998).

Such shifts in processing or skill mastery can be illustrated with the example that we used in this part as well: singing. For we discussed the fact that expertise in singing opera roles can result not only in learning different roles – like a Don Giovanni, a Wotan, and a Saint François – but also in learning to interpret these roles according to the requirements of a specific interpretation of the opera character. This can lead to a situation where an expert singer has interpreted a role in conflicting ways, requiring him to inhibit a stereotypical rendition of a virile Don Giovanni or serene Saint François in order to give way to an alternative interpretation. Interestingly, although much of our human behavior and cognition appears to be determined under circumstances by either automatic or controlled processing, we still have the capability of bringing about a shift from controlled to automatic processing. For the automatization of controlled processing we have several strategies available, such as those mentioned in the previous

²²⁴ As discussed in section II.3.1.5, our account differs from another account of skill acquisition that sharply distinguishes between expert intuition and the analytic reasoning that agents rely on before reaching that expert level (Dreyfus and Dreyfus 1986 ; Dreyfus 2004). In contrast, our account of kludge formation aims to explain how it is possible that experts are capable of fighting the capacity limitations associated with controlled processing in such a way that they remain capable of articulating or making explicit what they do. We don't want to deny that novices and experts share their reliance upon implicit knowledge, while intermediate experts will more often invoke explicit knowledge. However, it is quite possible for implicit knowledge to be made explicit, as when a neural network can offer a representation of its own specific state (Cleeremans 1997 ; Cleeremans, Timmermans et al. 2007). What we do deny, therefore, is that experts rely on a critically distinct process, different from those that novices use. Indeed, above we refer to an updated account of Simon's concept of chunking, that can explain why an expert may need to learn how to articulate his chunked and implicit form of information processing (Gobet and Chassy 2009).

section. These have observable results even in a situation where common controlled processing will not be viable due to pressure or the availability of limited cognitive resources. Apart from strategies aimed at a reduction of the information load of the upcoming process by segmenting or chunking the relevant information, an agent can also engage in self-regulation by employing explicit representations of the future situation or future response in order to prepare for a controlled yet quasi-automatic performance of the desirable behavior.

All such learning processes can be considered as dependent upon the formation of a kludge modifying the mechanism that can be held responsible for a singer's performance. That is, in all such cases the explanatory mechanism's activity leads to observable differences in its behavior, even though it may have remained largely the same as before. It may be that a kludge has formed such that keeping tone no longer requires ongoing attention, but is automatized. Or a kludge has been added to the vocal expression mechanism, enabling the singer now to reliably coordinate his singing with keeping rhythm. The kludge formations that we have observed so far appear as internal modifications, requiring no apparent resources from outside the agent. However, this may already be an oversimplification. In the next chapter we will discuss two types of kludge formation that obviously involve external resources.

After all, kludge formation that obtains during singing practice is not just an internal affair comparable to repeating a spontaneous behavior many times. Although our first kludge characteristic focuses on its functional properties, these may be shaped or sculpted under the influence of external or environmental information – which refers to our seventh kludge characteristic. Obviously, with this combination of characteristics, we can expect kludges to emerge in practically all domains of cognitive processing. An example that demonstrates the interaction between developing functional properties and environmental information is a baby's spontaneous babbling or crying which gradually gives way to more melodic vocal expressions. Contrary to the assumption that this process depends just upon repeating prespecified behavior, culturally specific external information about tonal mode turns out to play a role here from the moment that a baby starts to cry and increasingly so when specific scales are used for music and in singing.²²⁵ On top of these culturally specific influences, there are many other

²²⁵ Newborn infants cry differently in France and in Germany, reflecting the specific inflections of their respective mother tongues (Mampe, Friederici et al. 2009). As is well known, during their youth, children's capacities for the general recognition and production of specific vowels and consonants are increasingly constrained under the influence of exposure to their mother tongue(s). Similarly, the specific rhythms and scales prevalent in the environment will shape the space of a person's vocal expressions, even if he may later be able to expand or transgress that shape with effort. Indeed, young children have better memory of absolute pitch, probably because it allows them to better recognize the voice of their caregivers (Trehub and Hannon 2006). But even after early youth, the exposure to a pitch language influences absolute pitch recognition, explaining why Japanese children are better than Canadian children at memorizing pitch (Trehub, Schellenberg et al. 2008).

resources that an opera singer must learn to integrate in his performance. Consider, for example, the abundant use of external tools and information, like a music score, pictures, stage-properties and the like. Interestingly, performance is not always made more complicated by these external additives. On the contrary, performing a role can be facilitated by their use, as when Don Giovanni's or Saint François' particular singing lines become associated with the singer's successive and specific manipulations of a sword or cross, which are generally easier to remember. Even amateurs can experience that upon seeing a long forgotten score their voice automatically prepares for its first, high and dissonant, note.

It is to this phenomenon of integrating environmental information in the formation of a kludge that we now turn. Although it has already been mentioned earlier, it is of such importance that we will not close this part before treating it separately. It will confirm our earlier observation that our behavior and cognition can be considered to be the result of extremely 'open programs' (Mayr 1974) capable of integrating such information. Indeed, this integration can occur at such an early stage of development or learning that its result in turn is employed in subsequent behavioral or cognitive mechanisms. At first sight, such deeply and generatively entrenched kludges often look like innate mechanisms, although in fact they are acquired (Wimsatt 1986). In what follows we will discuss a few accounts of kludge formation with the involvement of external information. It will conclude our preparation for Part III, that will put forward a combined philosophical and cognitive scientific approach to a hierarchy of intentions.

4.1 Symbols, simulators and the malleability and stability of cognitive processing

With the two previous discussions on kludge formation, focusing on child development and learning, and on dual-process theories respectively, the impression may be that kludge formation comes naturally and only employs natural resources that are available without an important role for socio-cultural or environmental information. In order to ward off this impression, the present section will discuss theoretical arguments and evidence suggesting that humans are capable of developing 'simulators' such that these tie together a variety of content features employed in different conditions and for different purposes. This concurs with the general observation, which we referred to several times above, that evolving and developing mechanisms will usually benefit from the capability of integrating such environmental information (Wimsatt 1986 ; Wimsatt 2001). In what follows, we will observe that not only environmental information, but also internal information that is stored in a distributed distributed

way, can become integrated similarly in such mechanisms. This bestows upon such dynamic mechanisms several highly relevant properties.

Obviously, many dynamic mechanisms are capable of responding to environmental information on a momentary basis, allowing them to respond to environmental changes. However, in this chapter we are also interested in the integration of such information in a more enduring and stable form in mechanisms. If that happens, it will allow the further development of such a mechanism to build upon the integrated information, which will as a consequence become ever more entrenched into it. This malleability of the mechanism does not undermine its stability but instead enhances this, as the integrated information itself has contributed more directly to the mechanism's functional properties. As a consequence, the mechanism will generally respond more reliably and faster to related environmental information. The question that poses itself, however, is whether this development robs the mechanism of some of its adaptivity or flexibility, being less capable of responding to unexpected and novel environmental features. We will argue that a mechanism's adaptivity combined with its stability will depend upon the hierarchical structure of the information it has integrated. If it is capable of preserving a necessary hierarchy of this information, it may be able to process continuously novel environmental features with a largely unchanging and stable mechanism.²²⁶ The next sections will be devoted to the discussion of kludges that are constituted by a cognitive mechanism's interacting in an enduring and stable manner with specific environmental and internal information. Our discussion of this capability will first discuss a theoretical account of cognition that assigns an important role to cognitive simulation. Subsequently, we will discuss the concept of extended cognition, which implies that cognitive processes do not occur solely inside the skull but extend into the world, integrating in a tight manner not just environmental information but even objects and technologies.²²⁷

Although several simulation theories are on offer (Goldman 2006 ; Hesslow 2002 ; Zwaan 2008), we will focus on Barsalou's account, as it develops its argument such that it simultaneously allows a central role to symbols and language in human

²²⁶ The prevalence of theories that ascribe hierarchical structures to cognitive processes and representations is largely due to the assumption that such a structure offers several benefits, such as stability and speed in processing, multiple forms of access, and many combinatorial options (Cohen 2000)

²²⁷ The debate concerning the 'extended mind' has meanwhile been reformulated as 'the Hypothesis of Extended Cognition' (Clark 2011), which also better suits our vocabulary.

²²⁸ Zwaan focuses on language processing and offers a simulation account of language comprehension. Mental simulations are proposed for explaining the structures that can be noted in our comprehension of events or situations (Zwaan 2008). However, this account implies that sensorimotor activations in particular are involved in language comprehension, and is less interested in how language or symbols are stored in the brain than Barsalou's account is (Barsalou 1999c ; Zwaan 2009).

cognitive processing.²²⁸ We will start our discussion with this latter issue. Accounting both for abundant evidence of many forms of interactions or interferences between action, perception and speech that is visible in subjects' performances and for the development of apparently abstract concepts in human speech, Barsalou developed a theory of perceptual symbol systems (Barsalou 1999c). Though this may not be the only proposal for explaining those forms of interactions or for grounding concepts in multimodal bodily experience, it has two interesting aspects that merit specific attention.²²⁹ First, the theory carefully argues against the suggestion that sensorimotor experiences are stored holistically, in which case the correspondingly grounded concepts would presumably represent an object, event or action equally holistically. Instead, the theory of perceptual symbol systems defends the view that memories are stored in a distributed manner across the brain - not in a holistic manner but divided in many components or features. Indeed, dependent upon the subject's attention in a given situation on specific features of an object or event or action while neglecting other features, the composite yet structured symbol that is stored will comprise components different from those that would have been stored in another situation (Barsalou 1999c ; Barsalou 2009). Such a modal symbol is stored in long term memory in a distributed manner, where connections between perceptual, motor, cognitive and affective features of an object, event or action is formed by means of numerous associations (Barsalou 2008 ; Goldstone and Barsalou 1998). Based upon the strength of the associations between features, a result of their regular co-occurrence or their being included in the subject's focus of attention, these associated features are likely to be co-activated in subsequent situations.²³⁰ This brings us to the second aspect of this proposal.

The account denies that the retrieval or activation of memory of an object, event or action involves the reactivation of a holistic and faithful representation of it, as though it were stored comprehensively in memory. Rather, a memory activation requires the reactivation of a complex of several different features, stored at different locations in the brain and associated more or less strongly with each other. Barsalou calls this process

²²⁹ An earlier proposal for symbol grounding which also aimed to account for the presence of abstract and concrete symbols was presented in (Harnad 1990). Its original aim was more modest, as it did not seek to account for the interaction between cognitive functions in which symbols may play a role, nor for the top-down influence of symbols on perception, for example.

²³⁰ However, they need not be co-activated all the time, also allowing for the development of novel, amodal symbols. Indeed, even though abstract thought may have perceptual origins, it does not preclude the creation of abstract contents that comply to new rules (Goldstone and Barsalou 1998). Besides, abstract content may have an origin not just in perceptual and motor states, but also from internal states or introspection (Barsalou 2008). In any case, it would be problematic to derive all possible conceptual contents and structures directly from modal symbol systems, without any role for amodal and abstract symbols (Dove 2009).

simulation, about which he writes: “[a]ccumulating evidence suggests that simulation constitutes a central form of computation throughout diverse forms of cognition, where simulation is the re-enactment of perceptual, motor and introspective states acquired during experience with the world, body and mind” (Barsalou 2009 1281; cf. Barsalou 2008)).²³¹ Such re-enactment occurs in a multi-modal fashion, drawing together again many features that were stored in a distributed way following many experiences.²³²

The brain facilitates such comprehensive re-enactment with the ‘simulators’ that integrate the relevant contents, emerging as a result of repeated experiences. Subjects will develop such simulators in great numbers, depending on their attending to actions, introspections, objects, events, situations, etcetera. The activation of such a simulator will occur upon perceiving a relevant situation or upon using a relevant word or concept. It can involve not only the re-enactment of modal states but also of conceptualizations or motor states (Barsalou 2009). This is the reason why simulation appears to be ‘a central form of computation.’²³³

Importantly, such an activation or re-enactment of a simulator is always a dynamical process and will yield a visual, motor, affective or conceptual simulation that is partly dependent upon the specific situation the subject is in. As a result, a simulator activation in a given situation will yield simulations that differ between subjects, and may also vary from situation to situation for a single subject, depending partly on the strength of associations that have become integrated in the simulation (Barsalou 2002 ; Barsalou 2009). However, simulators are not stored or activated randomly. Interestingly, they do have a structure that renders them both flexible and with some stability: a hierarchical structure. It consists of a superordinate ‘frame’ level,

²³¹ Simulation gained much interest after the discovery of mirror neurons, which support the activation of highly similar neural networks under different conditions, for instance upon the observation, imagination, imitation, verbalization and performance of actions (Grèzes and Decety 2001 ; Hesslow 2012). It is hypothesized that mirror neurons are also implicated in the evolutionary development of language, gestures forming an evolutionary precursor to language. See e.g. (Arbib 2003a ; Arbib 2005 ; Jeannerod 2008 ; Rizzolatti and Arbib 1998). Barsalou recognizes this potential relevance of mirror neurons for simulation theories, yet raises the question why non-human primates show such different abilities even though they have mirror neurons at their disposal, too (Barsalou 2008).

²³² Partly because of the involvement of mirror neurons, simulation gets its experiential, multi-modal nature, making it a richer source of understanding other subjects than theorizing, according to Gallese & Goldman’s account of intersubjective understanding (Gallese and Goldman 1998). Nonetheless, most cases of mindreading are likely to depend upon a hybrid of both mirroring and theorizing, according to Goldman (Goldman 2006).

²³³ Concurring with this observation, Hesslow defends a “‘simulation’ theory of cognitive function”, arguing that a wide range of cognitive processes in fact converge on their being a form of simulation. Most cognitive processes are one of three forms of simulation: simulation of action, simulation of perception, or anticipation by way of simulation of action consequences (Hesslow 2002,242).

comparable to the category class, and a subordinate level, which includes contents referring to specific members of that category – such as specific cars that all fall under the frame ‘car’ (Barsalou 1999c, § 2.4).²³⁴ With this structure and with the possibility of combining the many components in indefinitely many ways, these simulators can also account for the productivity that characterizes language. What is more, simulators are also involved in the development of abstract concepts (Barsalou 1999c ; Barsalou 2008 ; Goldstone and Barsalou 1998).²³⁵

A final aspect of this account that merits our attention is the fact that as much as simulators may give rise to linguistic concepts, concepts in turn can activate and control the simulations that rely on such simulators (Barsalou 1999c). Instead of assuming a uni-directional influence of modal systems on conceptualizations, the account recognizes a bi-directional interaction in which rich experiential simulations can also result from a subject’s attending to a concept (Barsalou 2009 ; Barsalou, Cohen et al. 2005).²³⁶ Indeed, this concurs with Barsalou’s general assumption that language comprehension is not so much a matter of activating an archival memory but is better considered as preparing agents for situated action, “or at least to create the experience of situated action” (Barsalou 1999a 75).²³⁷

Experience and expertise play a crucial role in this account. Obviously, an agent’s simulation of a situated experience and his simulation of potentially adequate actions will be richer if he can rely on a rich history of many relevant experiences

²³⁴ The multilevel, hierarchical structure of a simulator is comparable to the account of chunks in (Gobet and Simon 1996), where a chunk consists of a template with several free slots that allow situation-dependent specification. Both theories refer to hierarchical structures and thus aim to account for easy and fast recognition of complex content while avoiding a holistic account of the activated, memorized content.

²³⁵ While referring to this structure and the consequent productivity and recursivity of language, Goldstone & Barsalou argue that instead of strictly distinguishing between perceptual and conceptual representations, and between perceptual and abstract contents, we should conceive these as continua (Goldstone and Barsalou 1998). First empirical neurophysiological evidence in support of this theoretical continuity came from TMS experiments, showing that subjects activate motor areas for processing both concrete and abstract concepts (Glenberg, Sato et al. 2008). Meanwhile, other lines of research have supported this evidence that sensorimotor areas are recruited for processing abstract concepts (Ghio and Tettamanti 2010 ; Lacey, Stilla et al. 2012 ; Pecher, Boot et al. 2011 ; Pulvermüller 2012). Barsalou’s theory is specific for its explanation of the combinatorial productivity of language and its offering a richer account of abstract concepts than merely considering these as impoverished concepts for concrete objects (Barsalou, Simmons et al. 2003).

²³⁶ In a different context, Goldstone notes that the handling of abstract mathematical operations has an impact upon subsequent perceptions made by a subject. It once more confirms the widespread phenomenon of neural re-use (Anderson 2010) of evolutionary early – perceptual - systems for later cognitive – mathematical – tasks (Goldstone, Landy et al. 2010). The simulators and simulations that are discussed in the present section are another example of such re-use.

²³⁷ Obviously, the idea that the brains of organisms store information in order to better prepare for future actions is not specific for this account. Such a notion of memory as a source of information for prediction is made more explicit in (Bar 2009) – which account also assigns an even larger role for top-down information or analogy seeking in the processing of perceptual information.

(Barsalou 1999a).²³⁸ Moreover, given enough experience, a simulator for a particular skill can become so entrenched that it may be activated automatically in a situation that appears to be relevant, without requiring conscious effort.²³⁹ Such simulator activation or re-enactment will also enable the subject to predict or anticipate how the situation will develop, as the simulator involves components that are not yet visible in the situation but that have become associated with previously experienced, similar situations: future actions, necessary instruments, likely personal feelings, etcetera (Barsalou 2009).²⁴⁰ Finally, and crucially, expertise will play out to the extent that experts will be more capable of exerting selective attention to relevant as opposed to irrelevant features of a situation or object, which influences subsequent processing of these (Barsalou 1999c).²⁴¹ As we already mentioned, a simulator is never a faithful re-enactment of a given object, action, situation and so on, but can draw together

²³⁸ This account clearly differs from Sperber's modularity thesis that ascribes 'massive modularity' to the human brain, claiming that it develops modules not just for perceptual functions, but also for many quite specific cognitive processes and contents. According to this thesis, even 'micro-modules' exist that have as their domain just a single concept (Sperber 1996). Applying the notion of modules so widely appears problematic, as it must subsequently explain why perceptual and conceptual processes interact or respond to similar features of a category, or why category mistakes occur – which is better accounted for by the simulators described in (Barsalou 1999c). Although there is no room here to expand on this, we also believe that the recombinatorial productivity that is visible in human concept use is easier to explain in terms of simulators than in terms of micro-modules.

²³⁹ As the simulators emerge as a result of recurrent associations, statistical processes in the brain are involved in several aspects of symbol and language processing, contributing to an ever more refined structure of those simulators (Barsalou 2008). Notwithstanding these statistical processes, there is a significant top-down contribution of specific attention or concept use that influences the activation of one or another simulator.

²⁴⁰ Similar to this account of simulators is the hypothesis that human concepts emerge from distributed activation patterns in sensory and (primary) motor areas following upon experience. A concept then functions as a 'cog' – a term that is again similar to our notion of a 'kludge' – in our brain, providing "general structuring for sensory-motor observation, action, and simulation" (Lakoff 2006 161). However, the authors take the embodied nature of concepts much stricter. For example, they assume that these cogs provide image-schemas like containment schemas, source-path-goal schemas, force dynamics schemas, orientation schemas and others grounded in bodily experiences. Moreover, the authors suggest quite a tight relation – including some isomorphism – between the role of a concept and the functional structures of the associated brain networks: "the inferential structure of concepts is a consequence of the network structure of the brain and its organization in terms of functional clusters" (Gallese and Lakoff 2005 468, in italics in original). They go much further than Barsalou in this, and don't seem to recognize a bi-directional influence between concepts and the simulations they may provoke. Finally, they attach a certain rigidity to the functional structure of cogs, which is not the case for Barsalou's simulators – which have characteristic 'open-endedness' and allow for multiple reconfigurations (Barsalou, Kyle Simmons et al. 2003) -, or for our kludges.

²⁴¹ Treisman correctly notes that attention may not be a 'unitary process', as it both precedes and contributes to our perception of a coherent scene, and as it may rather be the outcome of a competition between different objects or features for our attention (Treisman 1998). We may expect that expertise can have a specific role in this, as it supports the outcome of such a competition precisely due to biases that are a result of previously accumulated perceptual and cognitive experiences.

any possible configuration of their components – thanks in part to the flexibility of our attention (Barsalou, Simmons et al. 2003). Moreover, and this may be peculiar to humans, such attention may be directed jointly by two subjects, thus capable of coordinating their handling of a shared situation or object (Barsalou 2005). The latter case of joint attention will benefit when both subjects share largely the same kind of expertise, facilitating their attending to the same relevant versus irrelevant features.²⁴² Now that we have pointed out that expertise is reflected in the development of simulators and the subsequent ease with which simulations of objects, situations, experiences or actions emerge, let us close this section by drawing another comparison to our notion of kludge.

4.2 Simulators and the kludge characteristics

Do simulators according to Barsalou's account have properties that make them similar to kludges, in that they affect a mechanism, or activate it in a specific way such that certain cognitive or behavioral responses can be explained? Does this account acknowledge the impact of learning and experience such that its description of a simulator concurs with the seven characteristics that we attached to the notion of kludge?

The first and foremost characteristic of a kludge was that its functional properties, rather than other properties, demonstrated its emergence. At first sight, it may seem that the theory of perceptual symbol systems is more interested in its neural properties than functional ones, as Barsalou writes: “the basic definition of perceptual symbols resides at the neural level.” However, he continues with a clarification which shows why it is precisely from the functional property of conscious availability that we can infer a very general neural property of perceptual symbols: “*unconscious* neural representations – *not conscious* mental images – constitute the core content of perceptual symbols” (Barsalou 1999c 583, italics added). Indeed, even though his account ascribes a central role to certain - multimodal - processes that underlie simulators and simulations, it primarily aims to account for many functional properties of human cognition and behavior. For its main theoretical purport is to contest the division between cognitive functions like perception, cognition and action. It does so by criticizing an amodal account of distinct concepts or categories and by building a theory about how subjects' concept use depends upon multimodal simulators that enable them to prepare for

²⁴² Obviously, however, in case attention needs to be drawn to a previously unobserved feature, expertise may have an impeding effect. Nonetheless, given that expertise results in more refined and structured simulators, Barsalou's account suggests that even in such cases expertise may support the discovery and interpretation of such novel features.

situated action (Barsalou 2002 ; Barsalou 1999a). Not surprisingly, therefore, such preparation for situated action is evident from functional properties produced by the simulators, primarily in modified behavioral responses resulting from the development of such simulators. Similarly, they appear to affect sensory perception, as research has shown. For lesion studies suggest the involvement of particular neural areas in simulator development, but they do so by providing evidence that consequences of lesions are observable in patients' affected behavior or cognition, thus supporting the focus on functional properties of simulators (Barsalou 2009 ; Barsalou 2008).

Our second kludge characteristic implied that we cannot derive directly from a kludge's functional properties what algorithmic theory can describe its operation. Now Barsalou appears to have committed himself to a particular algorithmic theory - in Marr's sense (Marr 1982) - in which simulation plays a central role in different forms of information processing ranging from perception, speech, and action to introspection, as we learnt in the previous section (Barsalou 2008 ; Barsalou 2009). However, the specific form of the process involved in a particular simulation can be highly varied. A simulator refers to a 'distributed multi-modal system' that integrates an increasing amount of information as diverse as properties, relations, events and mental states that are related to a particular category, like cat or bicycle (Barsalou 2009). Such simulators are developed for: "any component of experience (or configuration of components) processed repeatedly by attention" (Barsalou, Simmons et al. 2003 89), including not just attended objects or a situation external to the subject but also those that are present during introspection. As a result of these processes underlying a simulator, it is indeed unlikely that we can derive from a particular simulator's functional properties what representations have been employed nor how these have been use. Think of an expert singer who can prepare his role by either drawing upon a specific experienced situation, or upon an idiosyncratic configuration of some features of a typical situation, or upon his repeated imagination of a certain opera scene.

As for the neural implementation of the 'distributed multi-modal systems' that underlie simulators, it is to be expected that this third characteristic cannot be very specific.²⁴³ Nonetheless, a general proposal has been made in (Simmons and Barsalou 2003), where the authors distinguish between at least two different components of the development of simulators and their neural implementations. To begin with, given that simulators are developed from situated experiences, sensory-motor

²⁴³ Barsalou acknowledges how its distributed nature also plays out in the equally distributed, underlying neural systems: "a situated conceptualization is a multi-modal simulation of a multi-component situation, with each modal component simulated in the respective neural system" (Barsalou 2009 1283).

systems are intrinsically involved. Experimental neuroimaging and lesion studies do indeed demonstrate that specific auditory, visual, and motor systems are activated when subjects process words related to specific categories or concepts (Simmons and Barsalou 2003).²⁴⁴ But a simulator is not just the re-enactment of an actually experienced sensori-motor activity, as a simulator typically also integrates information stored in memory and derived from other experiences or simulators that share a feature property with the activated simulator. Neural association areas are suggested as the neural implementation for such sharing of features among simulators, comparable to the hidden layers in a connectionist feed-forward network (Barsalou 2003). The idea behind this hypothesis is that conjunctive neurons reactivate the components of other distributed networks, underlying other simulators, that are related to a currently activated simulator (Simmons and Barsalou 2003).²⁴⁵ Other neural areas can be recruited as well to support a simulator, for example when introspective states are employed for the simulation of particular features or contents (Barsalou 2003) or when imagination creates novel and unrealistic simulations (Barsalou 1999c). In sum, neural implementations of simulators differ largely and may even differ from time to time when a specific simulator is activated, depending again on the selective attention or context in which it is activated.

That brings us to the fourth kludge characteristic, which hardly needs specific attention. As a simulator is continuously developing and as context and attention determine the specific informational contents that are being activated in a given situation, its variability and flexibility is large. This does not imply that each individual has a completely random set of simulators with idiosyncratic informational properties, fortunately. Given the important role of modal and embodied processing and the influence of statistical processing of environmental information for the development of simulators, many simulators and their properties will largely remain stable for a subject and will to a great extent be similar among subjects (Barsalou 1999c; Barsalou, Simmons et al. 2003).²⁴⁶ Indeed, two subjects need not have developed identical simulators and can still produce strikingly similar simulations of a concept in a given

²⁴⁴ The involvement of sensori-motor systems in language processing meets with relatively broad agreement. See e.g. (Beilock 2009 ; Deacon 2006 ; Jirak, Menz et al. 2010 ; Pulvermüller and Fadiga 2010 ; Tremblay and Small 2011).

²⁴⁵ The authors refer to Damasio's idea of 'convergence zones', which equally aims to explain cognitive processing of language and meaning as the product of distributed networks, with an important role for various types of convergence zones, responsible for the abundance of associations observable in language and meaning processing (Damasio 1989). A recent review on embodiment and semantic representation concludes that in addition to the crucial role of sensori-motor areas, different theories are 'converging on convergence zones'. The reviewers point out that on the basis of the literature it is plausible that processing of abstract concepts does not necessarily engage primary sensori-motor areas - as strong embodiment theories would have it (Meteyard, Cuadrado et al. 2012).

situation or text, for example when this text or situation offers enough constraints on their simulations (Barsalou 1999c ; Barsalou 2009). Obviously, the richer the amount of experience and the simulators developed by subjects are, the more likely it is that they can produce such comparable simulations.

From this last observation we can immediately derive that a simulator is ‘cobbled together’, as our fifth kludge characteristic states. We asserted repeatedly that simulations are not to be understood as the activation of stored holistic memories of objects, events or actions. Instead, the simulator involved stores many different yet related perceptual symbols, recruiting sensori-motor representations among others.²⁴⁷ These symbols are stored in hierarchically structured networks, with frames at upper levels and further lower level contents. The development and activation of such a structured simulator is dependent upon a subject’s attention, the use of concepts and their meaning, and so on (Barsalou 1999c ; Barsalou 2009). Critically, having stored many of such simulators, humans can do more than just reactivate them to simulate previously experienced objects, events and actions. On top of this, they are capable of developing novel, abstract concepts and of simulating nonpresent situations with each other, which may be a critical component of human evolution (Barsalou 1999b). In sum, a simulator can be considered as a kludge also in the sense that it is composed as a distributed yet structured network consisting of components that are involved in many other cognitive processes as well. Notwithstanding this composite nature, each simulator has a relatively stable – hierarchical – structure, which merits its being called a kludge.

Being cobbled together, does a simulator in turn play a role in subsequent developments or experiences and become increasingly entrenched, as our sixth kludge characteristic would predict? Barsalou argues that this is indeed the case and reviews evidence that shows how perception, imagination, speech and other processes are facilitated and accelerated by the presence of relevant simulators. Based upon such

²⁴⁶ Moreover, there is behavioral and imaging evidence that in a general sense a conceptual system and even its implementation are to some extent shared between species like monkeys and humans (Barsalou 2005). A more elaborate theory of the evolutionary development of human language from precursory systems shared with animals is developed in (Arbib 2011). That account assigns an important role to mirror neurons, which appear to be recruited both in sensori-motor and in other association systems. Mirror neurons are also assumed to underlie the activation of motor systems when subjects are processing sentences with abstract conceptual contents (Glenberg, Sato et al. 2008). See footnote 231 above on the relation of Barsalou’s account to mirror neuron research.

²⁴⁷ There are several lines of evidence for such an embodied account of symbol and language processing, for example when subjects recruit sensori-motor areas in fMRI experiments of language processing, or when patients suffering from lesioned sensori-motor areas have difficulty in specific language tasks. Such findings suggest that sensori-motor activations are not just epiphenomenal, but also functional in language processing (Barsalou 2008). Evidence also suggests a bi-directional influence between sensori-motor activations and language processing (Pulvermüller 2012 ; Pulvermüller, Hauk et al. 2005 ; Zwaan 2009).

simulators, a subject will implicitly anticipate and attend to specific information when perceiving an occluded scene or half-understood sentence, for example, which further expands these simulators and strengthens their role in future situations (Barsalou 1999c). Concurring with these arguments, one comment adds that knowledge structures like frames and scripts – which will figure together with schemas in our next part – are related to such simulators and similarly contribute to language comprehension and other functions in the form of background knowledge (Zwaan, Stanfield et al. 1999).²⁴⁸ Other support for Barsalou’s observation concerning entrenchment comes from research on implicit memory, which has shown how memorized items subsequently facilitate the processing of related items by several cognitive functions like perception and the imagination of the past and future (Schacter and Addis 2007c). Entrenchment of particular simulators can even be observed in religious beliefs and rituals, which often employ contents that are grounded in modality-specific brain systems. By doing so, religious practice contributes to the social dissemination and recurring activation of these simulators (Barsalou, Barbey et al. 2005). Interacting cognitive and socio-cultural processes are thus responsible for the generative entrenchment of a collection of simulators that grows with an agent’s expertise.²⁴⁹

Finally, for our seventh kludge characteristic we must consider whether environmental information is involved in the development and functioning of a simulator. Although simulators are grounded in perceptual symbol systems and differ in that respect from amodal accounts of symbols or concepts, we have meanwhile elucidated that this account emphasize the embodied nature of symbols and concepts as well as their situated nature. Environmental information does not just influence the development of simulators, often a simulator is activated in response to a situation (Barsalou 1999c). Experts, in particular, can be recognized by the fact that certain situations will automatically re-enact similar situations, with the activation of specific simulations and consequently facilitation of certain responses over others (Barsalou 2009). This results in the subject’s situated experience of ‘being there’, in stark contrast to a subject who can only try to apply amodal concepts to objects, events and actions that are different from the ones from which they were originally abstracted

²⁴⁸ As mentioned in a footnote 257 the evolution of language is considered to be a rather continuous process, with several species sharing many different language processing components with humans. Reviewing evidence about the language faculty in different species, the authors of (Hauser, Chomsky et al. 2002) suggest that it is communication in a very broad sense that drove this evolution, drawing upon communicative dispositions that differed from species to species.

²⁴⁹ Cultural evolution is in many ways dependent upon the generative entrenchment of specific contents and practices in the behavior and cognition of individual subjects. Humans in particular use many ‘scaffolding’ strategies that contribute to this process (Wimsatt and Griesemer 2007).

(Barsalou 2002). Finally, culturally specific information plays an important role in the development and activation of simulators, underlining the relevance of environmental information (Barsalou 1999c ; Barsalou, Cohen et al. 2005).

In sum, the simulators this section has been devoted to turn out to comply largely with our notion of 'kludge.' Given its wide range of application to many domains of cognitive processing, it offers strong support for our argument that the explanatory mechanisms responsible for our acting are extensively modifiable as a result of learning and experience. As a result, our space of actions is being sculpted, with some options gaining in probability to be performed, whereas other actions are less likely to come to light. Next, we will discuss a final source of support for our argument about the prominence of such modification, now focusing on our capability to develop kludges with the involvement of environmental information and even environmental objects or material structures.

4.3 Reaching outside the skull: how can external objects become integrated?

The first chapters of Part II were devoted to processes that appeared, at first sight at least, to be part of human natural development and experience. The phenomenon of modularization in infant learning and the development of automatized processing was not shown to rely on a specific and external source of information like language, even though it is likely that instruction does play a role in those cases. In the last section and the next, in contrast, we consider forms of kludge formation that more explicitly rely on the involvement of environmental information – and here of external objects. Simulators, which we found to play an important role in storing and employing such information, helped already to demonstrate that even culturally specific information will be integrated in cognitive mechanisms, influencing the functional properties of these simulators and the simulations depending on them (Barsalou 1999c). This is not surprising, as we already argued in chapter II.1 that we should not expect evolution to give rise to closed programs only, as these would only be beneficial for short-lived organisms with little time for learning (Mayr 1974). Organisms that live longer conversely benefit from their ability to entrench environmental information in mechanisms responsible for their behavior and cognition, with this entrenched information subsequently determining in part their development (Schank and Wimsatt 1986).

Such entrenchment of information during processes of kludge formation invites us to pause for a moment to reflect on the nature of cognitive processing. To the extent that cognitive processing involves external information, its results are always influenced by

it, obviously. Given Hebbian learning processes, these processes themselves will be modified according to the specific amounts and importance of different contents that they are offered for processing, including contents that derive from the environment: a vulnerable and quick rabbit will gain perceptual and motor expertise that is vastly different from the expertise of a human opera singer.²⁵⁰ When environmental information becomes entrenched with the formation of a kludge, it thus contributes favorably to the facilitation of future processes within a specific environment in comparison to processes that are triggered by a completely novel environment.

What should be emphasized, moreover, is that the external information that becomes entrenched in cognitive mechanisms can have a profound effect on the structure of subsequent processes. In addition to associations that have been gradually and mostly implicitly shaped between specific contents – like those determining the results of automatized processing according to the dual-process theories discussed earlier – cognitive processes can follow rules or structures derived from external information. In humans, this capacity of structuring and restructuring cognitive processes according to acquired, external information is strongly developed. Language and the use of symbols play a significant role in this capacity, even if this role is different from case to case. For example, apart from facilitating agents to restructure their space of actions, symbol use can also help to inhibit automatic, reward related responses to stimuli. A demonstration of this was given in research with chimpanzees. These had difficulty in maximizing their rewards in a reverse-reinforcement contingency task in which they would receive maximum reward only when they selected the unappealing minimum, and vice versa. They failed only when they saw the attractive candies involved, but not when these were represented by abstract numbers: then the chimps were capable of selecting the lower number, obtaining maximum rewards (Boysen, Berntson et al. 1996). Symbolic representation thus helped them to override a strong action tendency and it facilitated restructuring the relation between direct aim and the indirect outcome of their action.

Compared to animals, human capacity for language and symbol learning is significantly different and larger.²⁵¹ Indeed, this capacity is so crucial to human existence because our brains have co-evolved with language use and our prefrontal cortex has overdeveloped such that humans: “are not just adapted for symbol learning but for *fail-*

²⁵⁰ This is not to deny that there are also many innate predispositions or prespecifications at work that help determine the early learning outcomes of rabbits versus humans and their subsequent chances of gaining expertise in running versus singing. Prespecified biases play a role in the neuroconstructivist account of such expertise, seeking a middle ground between innately specified content-specific modularity and unconstrained ‘tabula rasa’ accounts of cognition (Karmiloff-Smith 1992 ; Mareschal, Johnson et al. 2007).

safe symbol learning” (Deacon 1997 415, italics in original). Meanwhile, humans have created an ecological niche strongly determined by this language capacity (van der Lecq 2012). Even though animals share some capacity for symbol learning, humans are particularly able to acquire and recompose symbolic structures as a means to prepare for novel environments and novel actions, instead of merely relying on the strength of previously exercised associations, the way animals do (Barsalou 2005). Indeed, acquiring insight in deeper formal structures facilitate the transfer of agents’ expertise in pattern learning from domain to domain, even when these are different such as letters and animals. Generally, it appears that even early perceptual processes learn to adapt to acquired formal rules and process visual stimuli according to those rules. For example, eye-tracking research shows that adept readers focus on multiplication problems before focusing on addition problems in mathematical formulas. Given this impact of mastered abstract mathematical rules and structures on human perceptual processes, these processes are said to be ‘rigged up’ with mathematical expertise (Goldstone, Landy et al. 2010).²⁵²

One could consider, as we do, these phenomena as proofs of the brain’s – and our human brain’s in particular – potential to integrate environmental information in such a way that it is not only deeply entrenched in its mechanisms, but also generatively involved in further developments and cognitive processes. Another view is to focus less on the integration of external information, but rather on its use or employment in creating novel and ‘hybrid’ cognitive processes, or in ‘extending’ the mind. Reasoning in that vein, Clark interprets such phenomena as a result of cognition being partly dependent upon a ‘symbolic environment’ which offers “additional fulcrums of attention, memory and control” (Clark 2006 300).²⁵³ This phrasing suggests, however,

²⁵¹ Even doves are capable of learning categories and simple categorization rules (Ashby and Ell 2001). Nonetheless, some continuity in symbol learning should not blind us for crucial differences in animal and human language capabilities. Whether it makes sense to distinguish between a broadly and narrowly conceived language capacity in order to differentiate between both capabilities requires a discussion beyond the confines of this dissertation (Hauser, Chomsky et al. 2002). What is more relevant to our purposes is the extent to which language and action processing are intertwined, which will be discussed more prominently in the next part. Interesting in that respect is the fact that Broca’s area is involved in processing not just complex structure in language, but also in action (Arbib and Bonaiuto 2007 ; Hamzei, Rijntjes et al. 2003 ; Koechlin and Jubault 2006 ; Nishitani, Schurmann et al. 2005 ; Willems and Hagoort 2007).

²⁵² Conversely, bodily gestures appear to scaffold our acquisition of mathematical cognitive operations, facilitating learning and influencing brain connectivity correspondingly. At later stages, gestures still play a modulating role during mathematical performances. De Cruz considers this to be supportive evidence for the extended mind hypothesis (De Cruz 2008). Dehaene, in his work on mathematical representations and reasoning, does not draw such consequences even though he acknowledges the transformative impact of linguistic – discrete - number representation on the (enculturated) brain (Dehaene 2001 ; Dehaene, Spelke et al. 1999).

that the environment remains external to those cognitive functions, while we have been emphasizing the integration of external information in cognitive mechanisms in such a way that it not just modulates but actually modifies those mechanisms and thus the cognitive processes they perform.²⁵⁴ For Clark, however, the relative externality of this information is shared with the externality of tools and artifacts, which are likewise employed by the brain on a general basis. As we want to distinguish the current discussion from the previous one about symbols and simulators, and given that external material objects cannot be integrated in these mechanisms in the same way as language and symbols, we will presently focus on the question how we should interpret the hypothesis of ‘extended cognition’, particularly with regard to the interaction with these objects.

In their much discussed article on the ‘extended mind’, Clark and Chalmers do indeed scrutinize the interaction of humans with external objects like pens, notebooks, nautical slides, calculators. They propose to interpret this interaction as a case in which an external system becomes coupled in such a reliable and robust way to cognitive processes going on in the brain, that such objects significantly expand and alter these processes. Famous is the example of Otto, who must rely on an external memory in the form of a notebook that he always carries and uses. Otto is dependent upon this external memory in a way comparable to a normal person’s reliance on his biological memory. If this is indeed accepted as a valid comparison, the authors claim that we may need to seriously reconsider our understanding of cognition and the boundaries of the system underlying cognition and behavior (Clark and Chalmers 1998).²⁵⁵ While traditional accounts of cognitive processing tacitly maintain the skull as a fixed

²⁵³ Clark refers critically to Dennett’s idea of ‘simulating’ a ‘more-or-less serial virtual machine’ by the parallel hardware of the human brain with the help of language (Dennett 1993, 218 ff.). In a 1997 paper Dennett discusses some insights of Clark’s (Clark 1997), yet maintains that language is a tool – yet a far-reaching tool – for thinking as it can influence the projects, rules, policies and so on of our thinking (Dennett 2000). Clark assigns an important role to language as an external prop complementary to the brain, thus leaving cognitive processes largely unaffected by language (Clark 1997 ; Clark 2008). In Clark’s version, therefore, one could find a ‘fixed properties’ view of the brain (Kirchhoff 2012), quite different from our view of modifiable dynamical cognitive mechanisms. However, Clark increasingly appears to acknowledge the dynamic properties of the brain and their importance for extended cognition, emphasizing now that: “the biological brain adapts, selects, and alters, its own internal routines” for optimal exploiting external resources (Clark 2011 459). Indeed, in addition he contends that “the biological brain is the essential core element”, suggesting a more moderate take on extended cognition.

²⁵⁴ Clark refers to Barsalou’s perceptual symbol systems (Barsalou 1999c), yet still stresses the issue of complementarity of such systems to the cognitive processes going on in the brain instead of recognizing how these symbol systems modify cognitive processing. He maintains that “language need not profoundly reorganize the shape and texture of the neural coding routines themselves” (Clark 2006 302), whereas Barsalou suggests that language – with its productivity, systematicity and capacity for recursivity – does have an impact on coding and re-activation of symbols (Barsalou 1999c).

²⁵⁵ Indeed, according to that view, the mind is a plastic, open-ended system, “fully capable of including nonbiological props and aids as quite literally parts of [itself]” (Clark 2003 10).

boundary, it would then be more appropriate to refer to ‘extended cognition’ carried out by an ‘extended mind’ (Clark and Chalmers 1998).²⁵⁶

Our capability to integrate not only external information but also objects and technology in our cognitive and behavioral routines has been emphasized by Clark and others variously and in sometimes rather poetic statements.²⁵⁷ Opposing what he considers outdated accounts of cognition that focus on our neural apparatus and tend to overlook or underestimate this capability, Clark contends that we humans: “make ourselves into new kinds of cognitive engine by (amongst other things) annexing and co-opting elements of external cognitive scaffolding as proper parts of hybrid computational routines” (Clark 2006 299).²⁵⁸ Pens and notebooks were relatively simple, but we also see more recent examples of information processing and communication devices that we take for granted even though they vastly modify and expand the cognitive tasks we are able to perform, therefore deserving to be called ‘pseudo-neural’ (Clark 2003 45). Still more invasive are appliances like an additional,

²⁵⁶ This argument is akin to Wilson’s hypothesis of ‘wide computation’, which objects to methodological individualism in psychology, as there are at times environmental features that play a role in an individual’s computational processes (Wilson 1994).

²⁵⁷ Or in Clark’s words, humans are potentially: “human-technology symbionts: thinking and reasoning systems whose minds and selves are spread across biological brain and nonbiological circuitry” (Clark 2003 3). Two objections have been made, both challenging the ontological status of the system that is assumed here. First, Clark’s view is said to fail in appreciating that cognition does process content derived from social practices or conventions, like humans do when they process linguistic or mathematical symbols, and so forth. A second objection concerns the causal processes underlying cognition, that are different from external causal processes – some of which might be related to cognitive processes. Both objections are presented as grounds for rejecting the extension of mental or cognitive processes with causal processes of a different nature some of which involve derived contents (Adams and Aizawa 2001). Now the first objection is related to our seventh kludge characteristic, stating that a kludge can be established with the integration of environmental information. Derived content, can indeed become integrated in cognitive mechanisms, as is demonstrated when language switching in bilingual subjects implies alterations of their cognitive processes and the underlying neural networks (Garbin, Sanjuan et al. 2010 ; Kobayashi, Glover et al. 2008). However, the extended cognition hypothesis does not deny such diversity, as the hypothesis in fact emphasizes that hybridity regarding content and causality characterizes many of the cognitive processes that we usually accept as instances of cognition (Clark 2008). We would rather raise another critical remark. Once we recognize that cognitive processes are capable of kludge formation in which process environmental information and objects can play an important role, we still are not forced to include the latter into an account of the cognitive mechanism or system itself. Indeed, it seems to us that the extended view still needs to argue why ontological notions such as we find in Aristotle’s *Metaphysics* (cf. the mereology given in V, 6), like coherence in time and space and the relative independent existence, are not applicable to the embodied cognitive system. To conflate two systems that are distinct in some basic ontological senses, even if tightly coupled, into a single system requires stronger arguments than those that merely show how each system by itself is already made up of different components. See for related mereological issues with regard to arguments in cognitive neuroscience our (Keestra and Cowley 2009 ; Keestra and Cowley 2011).

²⁵⁸ In his critique of the extended view of the mind, Sterelny also refers to the process of external scaffolding, but considers this merely as another kind of ‘niche construction’, as we can witness many evolving species doing likewise by building nests or laying scent trails (Sterelny 2010). This account is more in agreement with ours, as it still focuses on cognitive mechanism’s capabilities for kludge formation with the integration of environmental information but does not extend the mechanism itself out into the environment.

third – prosthetic - arm, a cochlear implant which allows ultrasonic sound processing, or an implanted retinal display directly connected to an external information database for processing the perceived environment, as these devices not only substitute our common capabilities but even add novel options for interaction with the environment. With these devices in place, we are allegedly no longer just conducting hybrid computational processes, but have become truly ‘biotechnological hybrids’ or ‘soft-selves’: “continuously open to change and driven to leak through the confines of skin and skull, annexing more and more *nonbiological elements as aspects of the machinery of mind itself*” (Clark 2003 137, italics added).²⁵⁹

Although the openness and modifiability of cognitive mechanisms has been at stake in this part all along, let us consider whether we can explain the interaction with objects, artifacts and technologies within the present framework, which emphasizes the formation of kludges with the integration of external information. Earlier, we noted that handling a tool may become integrated in one’s performance, as when an expert singer finds that singing a difficult Don Giovanni or Saint François melody is facilitated by his manipulating a sword or cross. So yes, external objects can even shape cognitive processes, suggesting that these objects somehow play a role in complex functional mechanisms.

A relatively simple example of the modulatory influence of objects can be seen in monkeys, when the receptive field of visuo-somatosensory neurons expands after a limited period of handling an arm-extending rake, the tool apparently being integrated into its body schema (Iriki, Tanaka et al. 1996).²⁶⁰ Further investigations supported this notion of the brain’s capability to flexibly integrate external tools in the activation patterns of sensorimotor areas that were previously determined by the confines of the body (Mahon, Schwarzbach et al. 2010 ; Peeters, Simone et al. 2009).²⁶¹ However, this phenomenon is still of limited value regarding the more far-reaching conclusions

²⁵⁹ Elsewhere, Clark identifies the body with the roles it plays in intelligent behavior, it being: “the locus of willed action, the point of sensorimotor confluence, the gateway to intelligent offloading, and the stable (though not permanently fixed) platform whose features and relations can be relied upon (without being represented) in the computations underlying some intelligent performances” (Clark 2008 207). The rejection of representation in this context is based upon a rather strict notion of that disputed concept. A more liberal interpretation of representation would allow for a fruitful employment of that notion in mechanistic explanations – see the discussion referred to in note 99 in chapter I.5 and elsewhere.

²⁶⁰ The authors of (Iriki, Tanaka et al. 1996) interchangeably use the terms of body schema and body image when referring to this phenomenon of neural plasticity. However, to concur with Gallagher, it makes sense to use body schema when referring to: “a system of sensory-motor capacities that function without awareness or the necessity of perceptual monitoring” (Gallagher 2005 24), while awareness, attitudes and beliefs play a role in one’s body image only.

²⁶¹ An issue that is yet to be clarified is whether birds or non-human mammals can create or discover new tools by employing some knowledge of physics, but to date there is little evidence for such sophistication in animals (Emery and Clayton 2009).

discussed here.

For the simple conclusion would be that tools like rakes and sticks merely co-opt the already existing neural representations of our body parts which are then flexibly adjusted – comparable to the reverse phenomenon when a neural representation of a body part in the brain is adjusted upon cutting a nerve to it (Buonomano and Merzenich 1998). The arm extending tools appear to modify the explanatory mechanisms involved in reaching, touching or perceiving only superficially, by adjusting relevant sensorimotor parameters.²⁶² Several anatomical studies have shown how as a consequence of tool use the areas of both somato-sensory neurons and visual neurons are enlarged, in this way extending the peri-personal space of the monkeys. PET scan research in humans, using not only rakes for reaching but also tongs for picking up objects, shows comparably enlarged activation sites (Maravita and Iriki 2004).²⁶³

However, tool use with a larger impact than just dimensional adjustments has been studied as well, potentially offering a more convincing demonstration of our being ‘human-technology symbionts’ in Clark’s sense (Clark 2003).²⁶⁴ With more complex tools, novel actions are made possible involving action means or goals that are impossible to obtain with mere bodily movements or their elongated versions. Experiments with monkeys and humans who learned to use both normal and reverse pliers (that allow grasping an object when opening instead of closing the handles)

²⁶² It is possible to disentangle the sensorimotor adaptations that occur with tool-use and the modification of the feedback that is provided by the ‘distalized’ visual target (Arbib, Bonaiuto et al. 2009). In similar research with rubber hands, subjects report illusory sensations of being touched when they perceive how a prosthetic hand or rubber hand is being touched, if these are carefully positioned and manipulated by an experimenter relative to their body or attached to it (White, Davies et al. 2010). This fact demonstrates that these objects are not merely used as external instruments but indeed are integrated in a multimodal body schema. Interestingly, patients suffering from schizophrenia are more susceptible to such illusions, interpreted as failures in processing body ownership. Such failures are relatively easy to explain, as they rely on failed detection of the incongruence between visual, tactile and proprioceptive signals (Jeannerod and Pacherie 2004 ; Thakkar, Nichols et al. 2011).

²⁶³ Holmes and others argue that there may at least be an additional role for attention allocation in the explanation of the phenomena reported about tool use and the extension of an agent’s peri-personal space, as that phenomenon seems to be vulnerable to interferences with attention (Holmes, Calvert et al. 2004). As such acquired neural extensions are not as stable as the projections that stem from body parts, one could indeed expect these neural extensions to be more vulnerable, indeed. Nonetheless, investigation of the changes in visual-tactile representations along the dimension that became extended with tool use shows that these changes occur in a gradient along that axis, ruling out that it is only a matter of attention allocation (Bonifazi, Farne et al. 2007).

²⁶⁴ An indication of the importance of the function of a tool rather than its mere dimensions in ‘sculpting’ the agent’s peri-personal space is provided by an experiment where a 60-cm. stick is used, with a handle positioned halfway at 30 centimeters. The extra, yet useless, 30 centimeters did not extend the peri-personal space, confirming other research that demonstrated the importance of the agents’ active involvement with a tool’s functions for neural and cognitive adaptation to the tool (Farne, Iriki et al. 2005). Indeed, only when agents intend to use the tool for reaching an object, their peri-personal space was found to become actually extended. The authors point out that both ability and intention to perform an act modulate perception in these cases (Witt, Proffitt et al. 2005).

show that movements and action goals are differentially coded in the brain, allowing flexible configurations (Cattaneo, Caruana et al. 2009 ; Umiltà, Escola et al. 2008).²⁶⁵In that sense, such novel actions would be better comparable to the novel situation simulations that are made possible by the configuration of endless, new and abstracted combinations within perceptual symbol systems (Barsalou 1999c), as we discovered in the sections above.²⁶⁶ Indeed, the interaction with tools or artifacts has been shown to: “change the way the human brain perceives the size and configuration of our body parts” (Malafouris 2010).

Now it is well known that premotor and parietal areas are activated not only by motor engagement with objects and tools, but also by the observation, the imagination or planning of potential motor actions (Jeannerod and Frak 1999). Apparently, motor representations do not just represent the complete, stored experiences of complex motor actions. On the contrary, these motor representations are stored in such a way that they allow the explicit or tacit recomposition of such actions in order to facilitate the prediction of potential outcomes when interacting with the environment (Jeannerod 2006).

Adding another layer of complexity, the brain is usually composing several potential action representations in parallel, requiring a selection of the single action that is eventually performed. All in all a complex task, involving a highly distributed network of parietal, motor and also prefrontal cortex areas (Cisek 2006). The integration of external objects and tools in potential actions does then require some extra – but not completely novel - adaptivity and flexibility, added to an actor’s expertise of tacitly composing action representations for bodily actions with familiar means and goals. Humans appear to be particularly good at this.

Indeed, studies of (over-)imitation in humans and animals that use ‘artificial fruit’, or opaque boxes that have to be opened via complex sequences of actions, demonstrate that humans are inclined and better able to observe, analyze and compose comprehensive representations in order to imitate complex actions and object use.²⁶⁷

²⁶⁵ Other research shows how handling a contemporary tool like a (manipulated) computer mouse similarly leads to novel representations in the brain. The authors demonstrated the presence in the cerebellum of multiple, modularized models of features of novel mouse actions, rather than models of the hand actions that were required for using the mouse (Imamizu, Kuroda et al. 2003).

²⁶⁶ The limitations in nonhuman primates in tool use may be related to their demonstrated inability to master hierarchical structure in language (Fitch and Hauser 2004), as both require complex computations and coordination at several levels of complexity. Study of gestures in gorillas show an equally limited repertoire with very little evidence of the invention of new gestures for novel situations or forms of idiosyncrasy (Genty, Breuer et al. 2009). Cognitive archaeology similarly suggests that tool use – primarily in the form of stone use and stone knapping – emerges well before the development of symbolic representations or of language (Malafouris 2010 ; Stout and Chaminade 2009).

They handle such representations at several hierarchical levels of specificity whereas animals tend to leave out intermediate steps and focus on overarching action goals only (Lyons, Young et al. 2007 ; Nielsen and Tomaselli 2010). Such differences have caused wonder, given the fact that there are clear homologues between monkeys and humans regarding their neural systems involved in reaching, grasping and manipulation of objects. This has led to speculation that perhaps monkey limitations in causal knowledge are responsible (Johnson-Frey 2003). Apart from that possibility, there is abundant evidence for the elaborate neural representation of actions in humans, which allows humans the elaboration of a much more extensive action hierarchy than animals (Badre 2008 ; Grafton and de C. Hamilton 2007). Such a representation of an action at several hierarchical levels enables a flexibility that also strongly supports the integration of familiar or novel tool properties– potentially including their causal properties and relations.

Associated with this flexibility is the cognitive task of discovering whether or not an object allows for a particular action, which requires the perception and representation both of an action goal associable with the object and of the behavioral steps required to reach that goal (Masson, Bub et al. 2011). More complex still are situations in which tool use involves adjusting a familiar object to a novel application, as was the case with the use of reversed pliers mentioned earlier, which was associated with modified action representations and underlying neural activations (Cattaneo, Caruana et al. 2009).²⁶⁸ In order to become the veritable ‘human-technology symbionts’ (Clark 2003) humans are said to be, it is important that humans gain the same sort of expertise in adapting and recomposing action representations with objects as they have for actions without objects.

Now it is predominantly the prefrontal cortex that supports the establishing and learning of associations between various perceptual and motor representations necessary for the configuration of complex actions involving external objects (Fuster 2000). Interestingly, comparative evidence suggests that the prefrontal cortex partly

²⁶⁷ This finding appears in contrast with evidence taken to demonstrate ‘rational imitation’ and teleological reasoning capacities in young children. They appear not to imitate an irregular use of the head for switching on a light if they can infer from the situation that it is equally permissible to use their hands (Gergely, Bekkering et al. 2002). However, in a review of overimitation research in which rational imitation is discussed as well, the authors point out that overimitation occurs primarily in cases of tool or artifact use, where the opacity of the means-ends relationships precludes such immediate conclusions about the (ir)relevance of the actions involved. Overimitation thus helps children to find their way in our ‘artifact-centric culture’ (Lyons, Damrosch et al. 2011).

²⁶⁸ Neuroimaging investigations of skilled performances with different tools - as with two computer mice with contrasting properties - show that the two skills activate different cerebellum locations, suggesting that different representations are formed for each skill (Johnson-Frey 2004). Given the large representations of skilled actions, there will also be overlapping components of these kindred representations.

evolved as an extension of the motor cortex, which would explain why it is so much involved in action control and why it enables increased capability of learning to compose and control complex actions, including tool use. Differences in tool use between chimpanzees and bonobos correlate, for example, with the size of their dorsolateral frontal areas (Stout 2010). Expertise in tool use is shown to have similar learning effects as expertise in common motor actions in an imaging experiment with humans who were observing tools being used in common and uncommon manners. Distinct effects were observable for expertise with goals and with the means of actions (Valyear, Gallivan et al. 2012). Such evaluation and selection between multiple action representations, implied in the adjustment of action representations, is supported by large and distributed representations of action features, which require prefrontal cortical activations in humans (Cisek and Kalaska 2010). Just like Barsalou's simulators are associated with complex and distributed representations, it appears that a similar type of representation underlies action and allows its composability and versatility.

Given such confirmations of the brain's capability to develop new action representations with the integration of new information about external objects, it does not come as a surprise that the ability to learn to control high-tech appliances can even be found in monkeys. Indeed, monkeys demonstrated fast learning to control a virtual grasping hand through brain-machine interaction. The brain-machine interaction involved both movement control with electrodes connected to the monkey's primary motor cortex and some tactile feedback by electrodes connected to its sensorimotor areas, not very different from normal tool use (O'Doherty, Lebedev et al. 2011). Such insights suggest that 'human-technology symbionts' (Clark 2003) probably differ more in degree than in kind from their 'monkey-technology' counterparts. Moreover, they confirm our notion that brains are generally capable of developing complex routines, taking into account relevant external information and properties of objects and artifacts.

So far, this part has focused on the dynamic nature of the complex mechanisms that underlie our cognition and behavior, highlighting the phenomenon of kludge formation which is partly responsible for shaping and reshaping the space of actions available to any agent – whether or not in symbiosis with technology. Let us now check whether kludge formation can account for some of the astonishing phenomena that the extended cognition hypothesis asks attention for.

4.4 Cognition-extensions and the kludge characteristics

The article that started the discussion of extended cognition emphasizes the functional properties of many cognitive extensions while defending the non-neural nature of these extensions, concurring with our first kludge characteristic which refers to it

being predominantly recognizable in functional properties. Consequently, it was defended to recognize any process as cognitive because of its functional properties, even if it is carried out by a part of the world and not inside the head (Clark and Chalmers 1998). This so-called Parity Principle is testimony of the fact that the extended cognition hypothesis stresses the functional continuity of processes going on in the embodied brain and in its environment. Indeed, Clark asks us to: “judge various potential cognitive extensions behind a kind of ‘veil of metabolic ignorance’” (Clark 2011 449), implying that differences in physical implementation of the cognitive (component) processes should not bear much weight. As is to be expected, this strict focus on the functional properties has been challenged as neglecting an adequate role for the physical system and its boundary conditions that would be responsible for these functional properties: these conditions are usually different for the brain and for some of its extensions or the technologies it uses (Rupert 2010). Indeed, as soon as we consider the involvement of external objects, artifacts, or technologies in the functional properties of extended cognition, problems arise. For example, it is difficult to see how the process of Representational Redescription that is often involved in development and learning (Karmiloff-Smith 1992) can have an impact not just on the cognitive processes involved in tool use, but also on the tools or cognitive extensions like notebooks and calculators themselves. It appears that Clark also recognizes that irrespective of its potential cognitive extensions, the main function of cognition remains the same, when he observes that the: “overarching goal of minimizing informational surprise can be served (...) by the canny longer-term structuring of an environment” (Clark 2011 454).²⁶⁹ Thanks to the brain’s capability of kludge formation while integrating in multiple ways the relevant properties of external objects and artifacts, we are capable of cognitively processing information that would otherwise have remained impossible. Even though it is now agreed that it is: “the biological brain [that] adapts, selects, and alters, its own internal routines so as more and more fluently to exploit the reliable presence of all those specific, culturally selected, tuned, and delivered, resources” (Clark 2011 459), these resources allegedly modified and expanded the brain’s cognitive capabilities.²⁷⁰

The second kludge characteristic focuses on the algorithmic theory that could

²⁶⁹ This observation is associated with the thesis that the brain is a ‘prediction machine’, continuously involved in anticipating future perceptions or actions based upon previous experiences (Clark 2013). Similarly, it has been argued that the brain aims to minimize prediction errors (Friston 2010), or is proactively anticipating the future on the basis of past and current experiences (Bar 2009). Common to these theses is the flexible involvement of complex, distributed representations, too.

account for the kludge formation – in the present context potentially including representations pertaining to cognitive extensions. Since it is an essential feature of extended cognition to be hybrid both in terms of the recruited resources and in terms of the information to be represented, the question is whether there is a particular algorithmic theory involved and if so, whether we can determine it. In the first account of extended cognition, it was language in particular that was considered to be the tool that has: “the major burden of the coupling between agents”, allowing us to “spread this burden into the world” (Clark and Chalmers 1998 18). If language plays such a central role, we might use linguistic knowledge to derive some very general features of an algorithmic theory that is associated with a particular kludge. But what happens when the hybridity of extended cognition extends to other options for the representation and computation of information, like when representations are used that are geared to our sensorimotor capacities, like levers, movements, and so on (Clark 2008)?²⁷¹ Intriguingly, it might be easier to develop algorithmic theories for such cognitive extensions than for language-dependent ones. For example, where it is principally impossible to reconstruct with certainty most cognitive strategies of persons living in antiquity, it may be easier to explain their cognitive extensions: their symbol systems and their arithmetic devices. In all cases we might expect that patterns of stability comparable to the stability that ensues upon kludge formation have been developed, but it is often easier that reconstruct the representations – implied in Marr’s algorithmic theories - describing the interactions with these objects are usually much constrained by the highly determined demands and affordances of these. For this reason, extended cognition often amounts to the emergence of ‘horizontally extended cognitive modularity’ (Wheeler and Clark 2008).²⁷² For example, the functionality of normal and of reverse pliers is easy to recognize from

²⁷⁰ An important question remains how we apply the Parity Principle. For example, it is difficult to see how Otto’s notebook would be continuously and effortlessly updated once new knowledge about the museum’s location or collection is obtained during an accidental discussion (cf. Clark and Chalmers 1998). In the biological brain, on the other hand, there is ever more evidence that the so-called ‘default mode network’, an identifiable network that becomes active in the absence of actual cognitive or motor demands, is precisely responsible for such maintaining and updating of information (Raichle, MacLeod et al. 2001 ; Raichle and Snyder 2007). Indeed, given that the integration of information is an important and complex cognitive task, such a network would play indeed a crucial role (Hohwy 2007). Updating external resources would require extra efforts, time and attention, which implies a transgression of the Parity Principle, it seems.

²⁷¹ Clark builds also on Wilson’s notion of ‘wide computation’, which involves external resources and alternative information structures like pen and paper and mathematical notations. Wilson focuses particularly on an explanation of computation and argues in that context for a non-individualistic account (Wilson 1994). Compared to that computational focus, Clark’s ambitions are much larger as they pertain to an analysis not just of computation, but of the human mind in general, as testified by his book titles “Natural-born cyborgs. Minds, technologies, and the future of human intelligence” (Clark 2003) or “Supersizing the Mind. Embodiment, action, and cognitive extension” (Clark 2008).

their appearance, each provoking specific neural activation patterns. Encoding of the specific movements afforded by the handles occurs separately from the encodings of the action goals enabled by the pliers (Cattaneo, Caruana et al. 2009 ; Umiltà, Escola et al. 2008). Obviously, not all tools are equally transparent and opaque artifacts or tools will not allow observers or users to infer their functionality and handling (Lyons, Damrosch et al. 2011). Still, it is probable that at some level of specificity we may determine the algorithmic theory of extended cognition even better than for cognitive processes without such extensions.

At times, we may be able to derive an algorithmic theory of a kludge somewhat easier in cases of tool use than in previously discussed cases of kludge formation. Does this also hold for our third characteristic, concerning the neural implementation theory? Originally, the hypothesis of operations performed by the extended mind where considered to be the result of two distinct yet coupled systems, to wit: the cognitive system and an environmental system or object (Clark and Chalmers 1998). However, over time the argument has emphasized consideration of a single yet complex system as the source of such operations, with a “complex cognitive economy spanning brain, body, and world” (Clark 2008 217). Obviously, that system will involve not just a neural implementation but also corresponding bodily and environmental implementations. An important difference with the kludges discussed in earlier chapters is the fact that the kludges involved in this economy tend to be less stable because of these complexes being ‘soft assembled’, making them “transient extended cognitive systems” (Clark 2008 158). Moreover, their presence is dependent on multiple and different kinds of conditions, like those on which the add-ons like instruments and pen and paper depend. Compared to the systemic integrity of the brain – which is not soft-assembled from situation to situation, integrating environmental objects that are available – such extended cognitive systems can be characterized by the presence of more sets of highly different constraints and limitations, related not only to the embodied brain but also to those objects and the interactions with these. As a result, such systems are generally much more vulnerable. Indeed, this lack of systemic integrity is for Rupert a reason to distinguish principally between extended cognition and non-extended cognition (Rupert 2009).²⁷² Similar but with a different emphasis, it is argued that the complex nature of these extended cognitive systems is characterized by more

²⁷² In their account Wheeler & Clark build on the neuroconstructivist accounts of Karmiloff-Smith (Karmiloff-Smith 1992) and others (Mareschal, Johnson et al. 2007), because these demonstrate that the emergence of functional modularity can occur during individual development and learning and is not just a result of evolutionary processes (Wheeler and Clark 2008).

than just a single translational input-output connection within the system. That is, whereas normal cognitive processes involve representations derived from physical sensory input, leading to physical motor output, in extended cognitive systems there are many more boundaries present where such transduction occurs. This is the case, for example, when cognitive processing relies partly on additional information representations, as when these are written down in a notebook, as in Otto's case (Weiskopf 2010). Indeed, as much as tight and stabilized interactions between cognitive processing and external information or objects is possible, considering the latter to be equally constitutive components of a cognitive system as components of the embodied brain seems an overstatement (Aizawa 2010).²⁷⁴ Even though extended cognition is comparable to some extent with kludge formation and particularly relies on the capacity for kludge formation, the implementation of an extended cognitive system differs significantly from the implementation of a kludge in the brain.

Differentiation between general kludge formation and the formation of kludges in interaction with cognitive extensions brings us to the fourth kludge characteristic, pertaining to the variation between stages of kludge formation. In our earlier discussions of variability during kludge formation, its stability and modifications over time were an important issue. In previous sections of this chapter we already dealt with kludge formation due to increased expertise with language and symbols, so we will here focus on the variations in the use of tools. In cases where environmental objects are used as tools – like the notebook and pen upon which Otto is dependent for much of his actions (Clark and Chalmers 1998) – this variability is not only dependent upon (embodied) brain processes, but is also dependent upon the wear and tear of those objects, on the weather conditions that may affect them, their replaceability, and so on. Independent of the level of expertise with particular tools,

²⁷³ The emphasis on the integrity of the cognitive system does not imply an underestimation of the role of external instruments or information for cognitive processing. On the contrary, Rupert argues: “the fundamental theoretical construct of virtually all successful cognitive science – whether computationalist, connectionist, or dynamicist— is that of a persisting architecture interacting with an ever-changing cast of external materials to produce intelligent behavior” (Rupert 2010 344). Such an observation regarding the involvement of external information or objects in cognitive processing is not new, of course. A comparison of the hypothesis of the extended mind with Hegel's idea of an objective spirit that supports individual cognitive processing demonstrates also some overlap in those analyses, for example (Crisafi and Gallagher 2010).

²⁷⁴ The determination of the boundaries of explanatory mechanisms depends upon the consideration of the ‘constitutive relevance’ of the component parts and processes involved. Craver appears to concur with the hypothesis of extended cognition that at times cognitive mechanisms ‘draw upon resources outside of the brain and outside of the body to such an extent’ that we may recognize these to be constitutive of these mechanisms (Craver 2007 141). Although Craver draws upon the notion of robustness as developed by Wimsatt (Wimsatt 2007), he is less careful in distinguishing between levels of robustness.

an extra variability is involved as soon as components with such different properties are assembled to form a kludge. The differential variability of the components of the extended cognitive system is particularly visible at the interfaces that we generally can localize in such a system, for example when an artist has great control over his hand, less so over the pencil in his hand, but probably less over the interaction between pen and paper (Rupert 2009 cf. p. 170). However, once we restrict our focus to those cognitive processes that are involved in tool use, we can observe again different levels of learning and expertise with corresponding changes in neural activation patterns (Valyear, Gallivan et al. 2012), suggesting that kludge formation is involved here, too. And indeed, mastery of a tool consists partly in the expertise with handling the tool's oddities and compensating for these. As a result of that, an expert tool user will demonstrate more stability in his performance with the tool than a novice will (Charness and Tuffiash 2008), again confirming the impact of the process of kludge formation.²⁷⁵ This does not erase, however, the fact that with extended cognition, agents have to cope with multiple sets of constraints that are valid for the various components of extended cognitive systems. As a result, we may observe more complex and differential patterns of variability, depending on the stability of the distinct components with their quite different properties.

From this diversity and variability of the components of kludges that emerge from extended cognition, there is a straight connection to the fifth kludge characteristic. When we mentioned earlier a kludge's 'cobbled together' nature (Clark 1987 291) it referred mainly to the neural and cognitive processes out of which a kludge is developed. In the case of extended cognition, we are looking at a wider range of resources. Obviously, this earlier notion is still valid in this context. Indeed, in Clark's early article on 'the kludge in the machine', he emphasizes that complex human cognitive processes are not completely novel but are grounded in and built from the "proto-cognitive capacities we share with lower animals" (Clark 1987 291). Gradual changes to a cognitive system with such capacities can have large snowball effects given enough time. Once one agrees to include into the system environmental resources as intrinsic components, the composite nature discussed here is even more obvious. Indeed, Wheeler & Clark argue that the human cognitive system is intrinsically open and is in fact: "a cognitive machine intrinsically geared to self-transformation, artefact-based expansion and a snowballing/bootstrapping process of computational and representational growth" (Wheeler and Clark 2008 3572).

²⁷⁵ Indeed, differences in activation patterns between experts and novices are also found when subjects are merely imaging the use tennis rackets, showing extended neural activations only in experts (Fourkas, Bonavolontà et al. 2008).

Above was mention of the capacity of the brain – of the expanded human brain in particular – to compose complex and flexible representations for interacting with the environment with the integration of external information in these representations (Cisek and Kalaska 2010). Instead of considering such interactions as demonstrations of: “the nonneural body and the nonbodily world as each capable of making key cognitive contributions” (Clark 2007 164), it is defended that this is just another example of niche construction, it being continuous with many other such niche construction phenomena in nature (Sterelny 2010). Such defense implies that calling these external resources ‘not alien but complementary’ to the brain’s contributions to extended cognition (Clark 1997 ; Clark 2008) is still underestimating the importance of the fact that these external resources are not produced by neurophysiological processes prevalent in the body and brain, making it relatively easy to distinguish them from neural components. Given the relative independence of the neural components of cognition from their cognitive extensions, it is not plausible to put the former on a par with the latter regarding their involvement in cognitive developments (Adams and Aizawa 2001 ; Rupert 2009).²⁷⁶ This is in contrast with the prevalent re-use which happen to neural structures (Anderson 2010), as was the case with modularized neural networks involved in child learning (Karmiloff-Smith 1992), with networks involved in automatized processes (Frankish 2010) and with simulators when they have emerged (Barsalou 1999c).

Discussing the sixth kludge characteristic seems less problematic than the previous ones, as it is part of the hypothesis of extended cognition that cases of ‘horizontally extended cognitive modularity’ (Wheeler and Clark 2008) or of hybrid thoughts or hybrid computational routines (Clark 2006) can become involved in further cognitive developments or trajectories. From his earlier analysis of ‘gradual holism and the historical snowball’ (Clark 1987) onwards, Clark has emphasized how earlier cognitive developments both open up and constrain the option space for later ones. What remains important in the present context of extended cognition, however, is

²⁷⁶ Responding to Rupert’s worries, Clark insists on the Parity Principle which focuses merely on the functional comparability of the bodily and environmental contributions to cognitive processes with those contributions that are going on in the head. Moreover, Clark concedes that it may be a difference in grain of their respective analyses that is partly responsible for their dispute, as the Parity Principle does not require ‘fine-grained identity of causal contribution’ (Clark 2007 168). In terms of mechanistic explanation, the analysis of this dispute looks different: after a first decomposition of a cognitive phenomenon, scientists will look for the mechanistic components that are responsible for the phenomenon or its phenomenal components. It may well be that in doing so, the lack of robust interactions or coherence between some component parts and operations or the presence of several distinct loci of control is such that it makes more sense to explain the (cognitive) phenomenon as the product of several – interacting, perhaps coupled – separate explanatory mechanisms. Mereological considerations matter here, again (see note 257).

the fact that not all components involved share the same scales of time and space for their development or are equally capable of development. Physical components like the eye or the swim bladder develop gradually during the course of evolution, affecting eventually whole species and their further evolutionary trajectories (Clark 1987). In contrast with those, it is more difficult to predict the 'epidemiology' of the cognitive representations or material artifacts involved in cultural developments (Sperber 1985). For example, cultural components of extended cognition may transform within only a single generation, but depending upon their representational format – in a specific language, for example – these form the environment of a community that is rather limited in its geographical distribution. As much as the authors describe extended cognition as a consequence of human nature's 'extensive openness to training and input-based modification' (Wheeler and Clark 2008), the authors unfortunately pay scarce attention to the differences in generative entrenchment between natural and cultural components – with objects presenting yet another class – and between their developmental trajectories. For when kludges are established, for example in cognitive mechanisms, and are then involved in further developments, as a consequence the previously established kludges are becoming deeper generatively entrenched in the organism. That is to say, it may be possible to differentiate between older or more foundational entrenched kludges and those of a more recent and superficial nature, building upon those older ones (Wimsatt 1986). Consequently, if the integrity of such a foundational kludge is being compromised this is likely to have a chain of effects, as when someone's language skill is compromised and his overall social and even cognitive functioning is disturbed. Applying this notion of generative entrenchment in order to differentiate between the various trajectories of kludges in biology and culture, might offer a highly welcome nuance to the hypothesis of extended cognition [cf. (Wimsatt 1999 ; Wimsatt 2001 ; Wimsatt 2006b ; Wimsatt and Griesemer 2007)]. It would again confirm Rupert's emphasis upon the lack of integrity of the system that underlies extended cognition by (Rupert 2009), given that there are great differences between the components of extended cognition in this respect. Even though comparative evidence, including archaeological evidence, suggests that tool making and the development of language capability have co-evolved in humans, human tools tend to change more rapidly and contingently than language does – while the brain's evolution occurs at an even lower pace (Stout and Chaminade 2009). So even when the two differ in their influence on brain evolution, language and tool use are comparable with regard to the complex and composed representations associated with these and which are differentially involved in subsequent developmental and cultural trajectories (Roepstorff 2008).²⁷⁷ Unfortunately, in his discussion of 'material symbols' Clark only

refers to linguistic and mathematical symbol use and shuns the question whether the material properties of tools can become equally integrated in human cognition and action only on the basis of their *cognitive* representation, which is what we are arguing here (Clark 2006). Nonetheless, our view appears to concur with Clark's more recent emphasis on the modification of the brain's 'internal routines' as the condition for exploiting external resources (Clark 2011). When an individual develops such routines in tool use and these give rise to kludge formation, they will alter the option space of his further developmental and learning options.

With that we arrive at the seventh kludge characteristic, referring to the involvement of external, environmental information in kludge formation. That external information is important for the hypothesis of extended cognition is no longer a surprise. Nonetheless, the term 'external' is somewhat difficult to define with respect to extended cognition, because cognition itself 'leaks' into the body and the world according to this hypothesis (Clark 2008). Setting aside the at times somewhat hyperbolic rhetoric, we've already noticed that the hypothesis does in the end not reject the perspective of cognitive science in putting the brain central to its considerations. Indeed, it is recently argued that this hypothesis concurs with a 'neurally-unifying predictive coding framework' according to which it is especially the brain's efforts to minimize informational surprise that unify all processes of extended cognition (Clark 2011).²⁷⁸ Still, environmental information and objects can and do play a crucial role in this task because of human nature's inherently 'extensive openness to training and input-based modification' (Wheeler and Clark 2008), as we learnt above. Constructing their own environments, including the objects that occupy these, humans are in fact constructing niches that maximally employ this openness to their modification at several levels of specificity (Sterelny 2010). The formation of kludges in the interaction with language and tools and the involvement of these kludges in subsequent developments leads to an ever greater entrenchment of external information in the individual's cognitive mechanisms, in socio-cultural structures and perhaps eventually in the evolving brain of humans generally.

²⁷⁷ In his review of comparative evidence concerning the development of ever more complex tools by humans, Ambrose does hypothesize that particularly the expansion of the frontopolar part of the frontal lobe in humans is driven by tool use. Basically, though, he emphasizes how language and tool require similar cognitive capacities for handling hierarchically structured representations: "Assembling techno-units in different configurations produces functionally different tools. This is formally analogous to grammatical language, because hierarchical assemblies of sounds produce meaningful phrases and sentences, and changing word order changes meaning" (Ambrose 2001 1750).

²⁷⁸ Several authors argue that the brain's main task is to engage in such predictive coding of anticipated inputs in its engagement with the environment (Friston 2005), as this task of the - 'proactive' - brain enables optimal interactions of the organism with that environment (Bar 2009).