



## 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*. 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, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

## 2 MODULARIZATION AS A PROCESS CORRESPONDING TO LEARNING AND COGNITIVE DEVELOPMENT\*

---

The reader may have been asking himself why we did not simply adopt the common notion of ‘module’, which largely overlaps our concept ‘kludge’. Indeed, the fact that we subscribe to the notion that explanatory mechanisms for cognitive functions – and many others as well – generally have a hierarchical and modular structure would support the adoption of the former. However, there are a number of reasons not to do so.

From its seminal discussion by Fodor in his 1983 publication of ‘The Modularity of Mind’ (Fodor 1983), the notion of module has received quite different interpretations. Fodor defended the presence of modular systems particularly involved in perception and language processing –at work in the periphery instead of the central systems of the brain, in his words. He suggested that such systems in the brain respond to some or all of the following conditions: the input system is domain specific, its operation is mandatory while allowing limited top-down access, it is fast, its operations do not rely on other sources of information, its output is rather shallow, it is performed by a ‘fixed neural architecture’ and recognizable by characteristic lesion patterns and finally it displays a characteristic developmental trajectory. The relevance of this notion is made clear with the following remark in the “Caveats and Conclusions” to his essay: “the limits of modularity are also likely to be the limits of what we are going to be able to understand about the mind” (Fodor 1983 126).<sup>151</sup>

Notwithstanding the relevance of the notion, there has not been general agreement regarding its definition or its empirical plausibility. A first obstacle has been that many authors interpreted Fodor in the sense that the complete list of conditions for modularity must be applicable for a particular system to be modular<sup>152</sup> – even though Fodor did not require this (Coltheart 1999). Consequently, authors could dispute the applicability of the notion in a specific case by merely referring to a single condition’s not being fulfilled in that case.

A second obstacle is that many authors associated the notion of ‘module’ with an additional condition, namely its innateness. This was not mentioned as a separate

---

\* On pages 371, 373, and 375, figures I, II, and III offer simplified representations related to the arguments made in Parts I, II, and III respectively. Fig. II is particularly relevant as a representation of the main contents of section II.2.

<sup>151</sup>This caveat concurs with the methodological emphasis that Simon and Marr put on a ‘near-decomposable’ (Simon 1962) or modular (Marr 1982) structure for a system to be explained – both authors being aware that it is unclear whether this requirement refers to an epistemological or a metaphysical demand, or perhaps both.

<sup>152</sup>As for instance happens in the relatively positive review article on the use of modularity and its future by (Thomas and Karmiloff-Smith 1998 246).

condition by Fodor, who referred only to the possibility of an innate specification of the information that a module is capable of handling (Fodor 1983). Nonetheless, this association with innateness led to the reproach that Fodor defended an ‘anti-constructivist nativism’, which would correspond with his underestimating the plasticity of the human brain and its influence on the formation of modules (Karmiloff-Smith 1994).<sup>153</sup>

A third obstacle has been Fodor’s denial that modules can be assembled: “in the sense of having been put together from some stock of more elementary subprocesses” (Fodor 1983 37). This would limit the number and kind of modules severely and restrict the applicability of the term largely. Arguing against this contention, Coltheart shows that Fodor’s own descriptions of particular language processing and visual information processing modules in fact contain references to modular component processes like a phonetic processor or a ‘form-concept dictionary’ (Coltheart 1999).

In more recent years, the notion has become not only widely used but also interpreted ever more loosely. Trying to capture the debate, a recent review suggests interpreting the different notions of modularity along five different dimensions used to determine a case of modularity: its physical structure, its cognitive functions, its information processing or computation, the information it employs and finally its development. The author’s attempt to find consensus in the debate about the most relevant dimensions – computational, informational and physical – seems implausible in light of the divergence he noted earlier (Seok 2006).

A more recent attempt at articulating modularity, together with an explication of its prominence in systems of various kinds, seems more promising to us. The notion of modularity has regained interest due to large scale brain imaging and computational studies that support the prominence and effectiveness of hierarchical and modular structures for the explanation of brain activation patterns and corresponding cognitive functions. Specifically with regard to the networks that underlie these activities and the topological structures of these networks, it appears that these structures are modular at several levels. That is to say: “many systems have the fractal property of hierarchical modularity, multi-scale modularity or “Russian-doll” modularity (Meunier, Lambiotte

---

<sup>153</sup> Generally, authors who defend the position that a module may be the result of some ontogenetic developmental or learning trajectory, have had difficulty with the notion of modularity. Reviewing most of the recent, relevant literature on the notion, Barrett & Kurzban conclude that it has been the “equation of modular with “fixed,” “innate,” and “static”” that has led to much confusion (Barrett and Kurzban 2006 642). Marcus also argues that both the notion of modularity and the notion of the innateness of cognitive functions are implausible and unproductive but are nonetheless often combined by authors. In contrast he relies on insights from recent approaches in the study of evolution & development that suggest that the brain is determined by both ‘prewiring’ and ‘rewiring’: “innateness is about the extent to which the brain is prewired, plasticity about the extent to which it can be rewired” (Marcus 2009 151)

et al. 2010 2). Reason for this prevalence is that such systems are rapidly and robustly assembled, according to the authors – demonstrating that their approach has implicitly removed the three obstacles that we mentioned. They note, however, that as much as modules can be assembled over time, ageing is also associated with changes in relevant hierarchical and modular structures and corresponding functional changes (Meunier, Lambiotte et al. 2010). With that they emphasize the dynamic nature of modularity, which will be part of our argument below. We will start with evidence from development cognitive neuroscience for the process of modularization and its involvement in the development of cognitive functions.

## 2.1 Neuroconstructivism and the relevance of modularization

Referring to ‘Fodor’s anti-constructivist nativism’ in her seminal developmental cognitive neuroscience publication “Beyond Modularity: A developmental perspective on cognitive science” (Karmiloff-Smith 1992), the author positions herself against a nativist view<sup>154</sup> that involves rejecting the possibility for a brain to function and learn at all if it was not born with some prespecified contents and processes already available in domain specific modules.<sup>155</sup> In contrast she and other authors defend a position which is aptly called ‘constructivism’ (Quartz 1999) – or ‘neuroconstructivism’. The latter position considers development and learning crucial processes that contribute not just to increasingly complex cognition but also to the ‘construction’ of an increasing complexity of the brain’s networks that includes their gaining in modular structures (Karmiloff-Smith 2009 ; Mareschal, Johnson et al. 2007 ; Westermann, Mareschal et al. 2007).

---

<sup>154</sup> Nativist positions with respect to modularity gained strength with contributions from evolutionary psychology as it presented evolutionary explanations for the brain as an evolved container of distinct expert systems (Cosmides and Tooby 1997) – leaving an all too modest role to individual development and thus inviting criticism that insists on modularization and a constructivist account of development (Wheeler and Clark 2008). Evolutionary psychology has been criticized among other reasons for its insistence on innate modularity – peripheral, massive or otherwise (Buller and Hardcastle 2000 ; Prinz 2006). Differentiating developmental, functional and mental modules while assigning a role to development in evolutionary psychology, Griffiths contends that the modules investigated in neuropsychology and in evolutionary psychology may differ (Griffiths 2007).

<sup>155</sup> To be precise, her neuroconstructivist position challenges both “Fodor’s anti-constructivist nativism and Piaget’s anti-nativist constructivism” (Karmiloff-Smith 1994 693) – leaving some role for innate constraints that partly determine further developments. This position is similar to the ‘open programs’ proposed earlier by Mayr (Mayr 1974) and further elaborated by Wimsatt (Wimsatt 1986). However, as mentioned above, Fodor did not require all conditions to be met for a case of modularity: “it is reasonably easy to think of psychological processes that are fast but not encapsulated, or involuntary but not innate, and so forth. The present contention, in any event, is relatively modest” (Fodor 1983 137, note 35). Nonetheless, Karmiloff-Smith overlooks this when she writes: “It is the co-occurrence of all these properties that constitutes, for Fodor, a module” (Karmiloff-Smith 2006b 10). The identification of Fodor’s position with ‘nativism’ is therefore exaggerated.

Observations of children learning skills in language, physics, mathematics and psychological reasoning are discussed in Karmiloff's 1992 plea for a 'developmental perspective on cognitive science'. Other skills, like the motor skills of playing the piano and solving Rubik's Cube, are treated from a similar perspective, even when adults rather than children are the learners here (Karmiloff-Smith 1992). That same developmental perspective was later applied to neurological disorders like the Williams syndrome or Specific Language Impairment, on the basis that such a perspective can better account for the widespread symptoms of these disorders than perspectives which assume that either a brain operates without modular structures at all or instead that a brain is innate with a modular structure that is similar for newborns and adults alike (Campos and Sotillo 2008 ; Karmiloff-Smith 1998 ; Karmiloff-Smith 2009 ; Karmiloff-Smith 2011 ; Karmiloff-Smith, Scerif et al. 2003 ; Mareschal, Johnson et al. 2007). As we are not specifically interested in disorders nor in infant development, we will limit our discussion here to those aspects of development and learning that appear to hold for adult learning, too. Then we will ascertain whether the seven kludge characteristics listed in the previous chapter, apply to it.

Observations show that development and learning consist of two distinct processes that to some extent even seem at odds with each other: proceduralization and explicitation.

Learning usually starts with the proceduralization of a particular skill, amounting to: "rendering behavior more automatic and less accessible" (Karmiloff-Smith 1992 17). This is complemented by its explicitation process, involving representations of the skill's domain. These representations contain increasingly "explicitly defined" components, which offer the child new opportunities for adjusting its performance, for example because: "the potential relationships between procedural components can then be marked and represented internally" (Karmiloff-Smith 1992 22).<sup>156</sup> According to the neuroconstructivist account, development and learning depend on 'Representational Redescription' of representations of the relevant domain of the skill, with the initial representations being implicit and the subsequent three stages involving increasingly explicit representations.<sup>157</sup>

The term 'modularization' refers to elements of these two processes insofar as they imply that: "input and output processing becomes less influenced by other

---

<sup>156</sup> This focus on representation has been maintained in a more recent neuroconstructivist account of development, distinguishing itself from a dynamic systems approach: "we consider representations as the central construct of cognitive development" (Mareschal, Johnson et al. 2007 89).

<sup>157</sup> This idea of a series of representations is akin to Marr's idea that visual information processing involves the processing of a series of successive representations of the retinal image until a 3 D representation has been formed (Marr 1982).

processes in the brain. This causes knowledge to become more encapsulated and less accessible to other systems” (Karmiloff-Smith 1992 15). Although modularization applies particularly to the proceduralization component of learning, there is a direct connection to the representational changes involved in such learning and the representations subsequently becoming available for explicitation, for correction, for their use in other skills, and other uses (Clark and Karmiloff-Smith 1993).<sup>158</sup> Indeed, this connection has been rendered in a more recent neuroconstructivist monograph as: “representations trapped within modules underwent an intrinsic process of abstraction even after behavioral mastery, which would eventually ‘offer up’ the knowledge within the module to other cognitive processes” (Mareschal, Johnson et al. 2007 213).<sup>159</sup>

These short descriptions of the processes involved in development and learning according to a neuroconstructivist approach may suffice for now, as our main goal is to consider whether this account concurs with our argument that the formation of kludges play an important role in development and learning generally.<sup>160</sup> Let us therefore consider the seven kludge characteristics mentioned in section II.1.1 and compare these with evidence stemming from research according to this approach.

## 2.2 Modularization and the seven kludge characteristics

The first issue pertaining to a kludge was that, although it is part of an explanatory mechanism, it must be characterized primarily functionally – its emergence should be visible in the cognitive or behavioral responses of an individual in a particular domain. The neuroconstructivist approach to cognitive development defended in (Karmiloff-

---

<sup>158</sup> The correspondence between increasingly separate and isolated processes with increasing need of specific processes aiming at the interaction of such separate processes is in a review called the combination of ‘dissociation’ and ‘integration’ in cognitive development – itself again associated with neural developments (Johnson and Munakata 2005).

<sup>159</sup> The authors identify the prefrontal cortex as likely involved in this process of interaction and mediation between specializing modules (Mareschal, Johnson et al. 2007). This is indeed a plausible candidate for such a role, given the fact that prefrontal cortex is highly connected to many long-distance neural areas involved in the processing of rules and several linguistic elements, but is not engaged by ‘routine, automatic and overlearned behavior’ (Fuster 2001).

<sup>160</sup> Comparable in these respects with the neuroconstructivist account of development and learning is the account of implicit learning developed by Cleeremans and colleagues. Its main argument is that implicit learning involves changes in the relevant representations that only gradually become accessible for conscious and verbal processing, with cognitive and behavioral improvements obtaining already at earlier stages. In this account, too, modularity is not prespecified at birth and a static phenomenon, but rather dynamic and emerging phenomenon, associated with learning (Cleeremans 1993; Cleeremans 1997; Cleeremans and Jiménez 2002). Indeed, reference is made to ‘functional modularity’, leaving undetermined whether a neural form of modularity is associated with this functional characterization of learning a cognitive or behavioral task. Similar to the neuroconstructivist account, this implicit learning account also assumes that learners possess and control an increasing number of different representations of a particular task, yielding them correspondingly more options for performing it either automatized but hardly conscious or conversely conscious but not as fast (Cleeremans 1997).

Smith 1992) is indeed primarily based upon observations of such functional changes, involving the observed processes of proceduralization and explicitation mentioned above. Although aiming primarily to account for the specific details of these observed processes, the approach also seeks to provide explanations in terms of the changing representations involved, and to suggest possible underlying neural processes. An example of such specific observable results is the fact that these parallel processes do not lead to steady improvements regarding both, as an increase in behavioral errors occurs during phase 2, due to the subject's disregarding some external information in comparison to earlier and later phases (Karmiloff-Smith 1992).<sup>161</sup>

The current modularization account concurs with the second kludge characteristic, referring to the impossibility of deriving a particular algorithmic theory from the functional characterization of a kludge or to the possibility of there being more than one algorithmic theory that can account for a specific case of kludge formation. Nonetheless, this account does assign a crucial role to information processing and representations, as was noted in the previous section where we discussed the process of explicitation, it being complementary to the process of proceduralization.<sup>162</sup> Both these processes allegedly rely on several phases of Representational Redescription. The first implicit and three consecutive explicit phases are characterized by different representations of the same cognitive or behavioral task, varying with regard to foregrounded or abstracted task aspects, being increasingly context-free, and so on. The redescriptions can help to explain the differences between phases in cognitive and behavioral responses like the degree of awareness or verbal rendering of these responses. Finally, the account contends that this Representational Redescription process occurs for each task domain separately and perhaps with different tempo and timing (Karmiloff-Smith 1992),<sup>163</sup> while elsewhere it is elaborated how this

---

<sup>161</sup> An explanation for behavioral inconsistencies – or temporary setbacks - associated with a gradual improvement of a task's representation by a child is provided in terms of a bifurcation stage, signaling a non-linear process leading to qualitatively new knowledge in (Raijmakers 2007). A correlation between neural dynamics and inconsistencies in rule learning has been found that is in accordance with this computational result (Durstewitz, Vittoz et al. 2010).

<sup>162</sup> The complementarity of these processes transpires from the observation that development can be considered as: "a progressive increase in the complexity of representations, with the consequence that new competences can develop based on earlier, simpler ones" (Sirois, Spratling et al. 2008 322). The authors distinguish their take on representations – focusing on their increasing complexity – as being different from the earlier neuroconstructivist account that is central in our present discussion. This account is more focused on the increasing abstractness of the novel and differently formatted representations, which should help explain their involvement in different cognitive and behavioral procedures (Karmiloff-Smith 1992).

<sup>163</sup> There may be spill-over effects between domains, as was found in bilingual children who demonstrate more flexibility in drawing non-existing objects than their monolingual peers do, which is thought to rely on the former children's' expertise with handling different linguistic representations of identical objects (Adi-Japha, Berberich-Artzi et al. 2010).

Representational Redescription process provides a plausible model of information processing that can help explain some recurrent results of development and learning, including temporary regress in proficiency of a certain task, initial difficulty for an expert to explain his performance, and integration of expertise with other tasks (Clark and Karmiloff-Smith 1993).<sup>164</sup> It can be concluded, therefore, that there may be more than one algorithmic theory involved in a particular case of kludge formation or in a performance that relies on such a kludge. Nonetheless, even though it is impossible to derive from a particular kludged performance which algorithmic theory or representation is involved, this does not rule out the possibility of influencing this kludge formation process or performance by choosing a particular representational format of a given task. Obviously, such a format choice will have wider consequences as it may impact upon associations with other tasks or processes, as when a particular music representation format is chosen. In our Part III, we will have more to say about this.<sup>165</sup>

Thirdly, we stated that a kludge is not definable in terms of its neural implementation, even though changes in a specific set of behavioral and cognitive responses will primarily give reason to assume that the mechanism responsible for these responses has been modified such that a newly emerged kludge is part of it. However, given the prevalence of hierarchical and modular structure in explanatory mechanisms, we can expect that learning does involve a modification of precisely this structure. Indeed, the original neuroconstructivist account already contained a concrete conjecture regarding possible neural correlates of such a modification: “if the modularization thesis is correct, activation levels should initially be relatively distributed across the brain, and only with time (and this could be a short or relatively long time during infancy) would specific circuits always be activated in response to domain-specific inputs” (Karmiloff-Smith

---

<sup>164</sup> See note 160 above for the implicit learning theory, which in a similar way refers to representational redescription for explanations (Cleeremans 1997). Modelling implicit and explicit learning in order to account for empirical data on these learning processes, Sun and others are also interested in interaction between the two processes and their quite distinct task representations. Eventually, they argue for interaction between connectionist models for implicit processes and symbolic models for explicit processes, accounting for cases of inflexibility in implicit learning and of ineffectiveness of explicit re-learning of a given task (Sun, Slusarz et al. 2005).

<sup>165</sup> Indeed, research shows that differences in practice structure do not only lead to differences in skill learning, but also in the underlying neural processes. This can be explained in terms of the subjects' efforts to distill useful representations of the task at hand. Subjects learning a skill via variable practices appeared to engage higher level planning of the task with the involvement of prefrontal cortex, while subjects that practiced by mere repetition did recruit motor cortex only with less flexible control of the task (Kantak, Sullivan et al. 2010). Ethological observation in learning and imitating primates suggests that limited capabilities of representation at several hierarchical levels of a task and its components suffice to explain these processes and the shortcomings of primates in comparison to humans (Byrne and Russon 1998).



1992 5).<sup>166</sup> Even though this conjecture has received ample confirmation, it still leaves room for different specific neural implementations of this modularization process, as there can be several modifications of the responsible mechanism that can account for it.<sup>167</sup> In sum, notwithstanding the fact that modularization or increasingly focal neural activation patterns appear to be prime candidates for the neural implementation of kludge formation, it is not possible to identify kludge formation generally with a specific neural implementation.<sup>168</sup>

Is there indeed variation visible between individuals during the period leading to kludge formation regarding a particular cognitive or behavioral task, as our fourth characteristic implies? That is, should we expect identical changes in information processing or identical neural correlates to accompany the kludge formation process in different individuals, or is there room for variation even if the end result is comparable? The neuroconstructivist account suggests that although the final phase of mastering a particular task involves the individual's capability of handling multiple representations of the task simultaneously, it does allow individual differences in mode and timing of developing this capability (Karmiloff-Smith 1992). These differences are inevitable when one considers development and learning as the outcome of the dynamic interactions between systems, leaving intact the possibility that these interactions result in relatively stable cognitive and behavioral capabilities that are supported by developing modules (Elman, Bates et al. 1997). Indeed, gradual changes in a developing connectionist network can lead to abrupt and distinctive changes in cognition and behavior, in agreement with the modularization hypothesis of the neuroconstructivist approach.<sup>169</sup> In sum, we cannot but expect there to be relevant

---

<sup>166</sup> Below, in our discussion of the sixth kludge characteristic from the perspective of the modularization, we will emphasize how this perspective has difficulty with the idea of innate, neurally specified modules, as it is then hard to explain widespread consequences beyond a task domain of failures in the relevant modularization process (Karmiloff-Smith 1998).

<sup>167</sup> Meanwhile, brain imaging investigations in developing and learning subjects have been carried out that confirm the expectation, that: "developmental changes in patterns of brain activity appear to involve a shift from diffuse to more focal activation, likely representing a fine-tuning of relevant neural systems" (Durston and Casey 2006 2154). This confirmation does not only apply to infants but also to adults, as learning generally has been shown to be associated with an increasingly modular – and hierarchical – structure of the responsible neural mechanism (Bassett, Wymbs et al. 2011).

<sup>168</sup> It has been critically remarked that a wholesale denial of innately specified processes in the infant brain by some neuroconstructivists is at odds with some – although limited – neuroscientific evidence for such processes (Franck 2004). However, neuroconstructivist approaches do assign a role to some innate domain-specific predispositions – like attention biases – and even to some innate specifications which can be triggered by environmental stimuli, but argue that it is further interactions with the environment that subsequently determine the cognitive processes (Karmiloff-Smith 1992 ; Mareschal, Johnson et al. 2007)

<sup>169</sup> The observation of such abrupt changes has been described and underpinned with a tentative explanation in (Karmiloff-Smith 1992). More recently, a computational account that refers to bifurcation stages in developing networks has been added to such an explanation (Raijmakers 2007).

differences between individuals, at least during the intermediate stages of the formation of a kludge. These inter-individual differences can pertain to all aspects of kludge formation that were mentioned in the previous three characteristics – overlapping largely with the computational, algorithmic and neural implementation theories distinguished by Marr (Marr 1982).<sup>170</sup> This prevalence of inter-individual differences holds even if developmental processes generally yield increasingly hierarchical and modular structures.<sup>171</sup>

The fifth kludge characteristic refers to the meaning of its name, as a kludge is not innate but derived from already present components – functions or processes, or a combination of both - with properties that are different from those of the eventual kludge. Here, the modularization account presents a nuanced answer that eventually supports the fact that re-use does not only occur with regard to neural areas but also to functional properties. Although modules allegedly are the result rather than the beginning of development and learning, these processes do not start from scratch in infants. Instead, a ‘skeletal outline’ of the brain that is present from the start, including several biases and predispositions, develops into an increasingly complex and modular structure, with changing connections to different brain processes (Karmiloff-Smith 1992 15, ff.). Nonetheless, an important difference exists between infant and adult brains, the latter containing more modularized structures than the former, corresponding with an increase in rather domain-specific cognitive processes (Karmiloff-Smith, Scerif et al. 2003). Associated with this is the fact that deficits existing at an early age, before extensive modularization has occurred, will have observable consequences in more than just a single cognitive domain because cognitive domains are only gradually isolated from each other: “the fact that domains are highly interrelated early in brain development (...) turns out to play a critical role in the formation of more general, albeit sometimes subtle, deficits in later development” (Karmiloff-Smith 2006a 47). In sum, modularization according to this account is indeed similar to kludge formation, modifying a mechanism that is responsible for a particular cognitive or behavioral task such that its weak modular structure with domain-general properties strengthens and its functional properties change, including its changing connectivity to mechanisms responsible for other functions.<sup>172</sup>

---

<sup>170</sup> Responding to the neuroconstructivist account presented in (Sirois, Spratling et al. 2008), Bateson adds to its interest in individual differences in development that it is not only the active role of the individual that matters but also its particular environment that influences development (Bateson 2008).

<sup>171</sup> Even though development generally builds on and further expands the hierarchical, modular structure of the brain, modifications of this structure depend so much upon interactions between different levels of the brain and its functions that a great deal of individual variability will be inevitable (Bassett and Gazzaniga 2011).

This latter aspect of changing connectivity brings us to the sixth characteristic: a kludge's involvement in further dynamic trajectories, its potentially being integrated as a component into a mechanism that underlies another function. For example, the observation that early visuospatial deficits in infants can have "cascading developmental effects over time on several *emerging* higher level linguistic and cognitive domains" (Karmiloff-Smith 2006a 47) confirms the occurrence of wide-spread neural 're-use' in the brain and the corresponding widespread consequences particularly of early functional deficits, as multiple functions will rely partly on shared neural components (Anderson 2010).<sup>173</sup> The question is whether such re-use also pertains to functions that have modularized due to development or learning, or whether modularization would preclude underlying mechanisms of such functions from being 're-used' again in other functions.<sup>174</sup>

Given Karmiloff-Smith's focus on developing brains and developmental disorders, this question is not a primary focus in her work, some of which concerns the wide-spread consequences of deficits in not-yet-modularized functions at later stages (Karmiloff-Smith 1998). And although the notion of modularization is attached to domain specificity, she does acknowledge that within any domain there are 'micro-domains' that function in turn as subcomponents – suggesting that micro-domain modules pertaining to gravity or pronoun acquisition can indeed be integrated in other, more comprehensive modules recruited for physics of language processing (Karmiloff-Smith 1992). Besides, although the prevalence of dissociations in adult brains suggests that modules are mostly domain-specific and segregated in adults, there are still domain-general disorders that rely on more focal deficits affecting a particular and modular function, as is the case with prefrontal deficits (Karmiloff-Smith 1992). Not just in the infant brain, therefore, but also in the adult brain we can observe that modularization can in fact facilitate the interaction of a modularized

---

<sup>172</sup> A more recent version of neuroconstructivism explicitly explains the development of complex representations on the basis of previously established components, relying partly on modularization processes (Mareschal, Johnson et al. 2007)

<sup>173</sup> As noted in footnote 100 in part I, neural re-use is a prevalent phenomenon in the evolution and development of the brain. As a consequence, it was emphasized in that context that it is necessary to employ 'domain-neutral' terms to refer to component mechanisms that are being used in multiple cognitive and behavioral functions (Anderson 2010). Both issues concur with aspects of the presently discussed modularization account. Indeed, Anderson himself has acknowledged agreement between his approach and the neuroconstructivist emphasis on modularization combined with interregional connectivity (Anderson 2008).

<sup>174</sup> It is worth noting that early connectionist models were 'highly task-specific and single-purpose', leaving unfulfilled Karmiloff-Smith's demand for developing models which go beyond modeling mastery of a specific task and that allow transfer of information across domains (Elman, Bates et al. 1997). It appears plausible that this limitation of these early models has contributed to rather limited interpretations of modularity, emphasizing informational encapsulation instead of allowing such information transfer, for example.

function with other functions, or its integration as a component in a more complex cognitive function.<sup>175</sup> Indeed, the twin processes of proceduralization and explicitation refer precisely to the increasing availability of a domain of expertise - which is being processed in a modularized way - to other processes. Thanks to the representational redescription which happens to such expertise during this process, its representational constraints render it available for both intra-domain and inter-domain relationships (Karmiloff-Smith 1992). Expertise, for example, can be held to rely on representational redescription yielding the availability of tacit or implicit knowledge to awareness (Feist 2013).<sup>176</sup> In Part III, we will discuss more explicitly how an agent with habituated intentional action patterns is better capable of flexibly responding to environmental information, as his action performance absorbs fewer resources than in the novice and leaves other resources available for processing this information and necessary action adjustments.<sup>177</sup> Several processes are involved in this development, of which the formation of relevant kludges is an important one, as we will see.

In this development of relevant kludges, the integration of environmental information plays an important role. The modularization account, in opposition to nativist accounts, not only emphasizes the importance of developmental and learning trajectories, but also the role of environmental information in these processes. It does acknowledge that an infant appears to be born with domain-specific predispositions

---

<sup>175</sup> Above we already noted that Fodor's notion of modularity should be supplemented with the notion of an assembly of modules, where component tasks rely on separate modules integrated in a complex assembly for the comprehensive task, as is defended in (Coltheart 1999).

<sup>176</sup> Expertise plays out differently in neural activations, as found in several fMRI experiments. Sometimes it appears that experts' brains show less activation in carrying out a task, sometimes experts recruit more neural areas. An explanation for this divergence in results could be that performance of a particular expert task must rely on the combination of a proceduralized and modularized component task with another cognitive task, while more focal activations suggest that this other task need only recruit a less complex mechanism. Indeed, not just a single form but various forms of the reconfiguration of - hierarchical and modular - neural networks are needed to account for both experimental and modeling results, depending among other things on task complexity (Bassett, Wymbs et al. 2011). Early imaging results suggested, for example, that an expert's skilled task performance allowed him to perform without much control and thus allowing him to again spend attention to information not directly relevant to the task, corresponding with increasingly widespread activations in comparison to novices (Raichle 1998). Imaging results for a motor sequence task in another experiment led to increasingly focal activation patterns for two distinct components of that task (synchronization and accuracy) and increasing connectivity between the neural mechanisms underlying these components (Steele and Penhune 2010). These examples show that it is highly task dependent how skill learning or expertise impacts on neural activation patterns and that such pattern changes can be multiple during the learning trajectory, but that in general such impact of learning is very common (Tracy, Flanders et al. 2003; van Mier, Tempel et al. 1998; Yin, Mulcare et al. 2009). Needless to say, the changing neural activation patterns eventually depend upon changes in synaptic activities at the single neuron level, which have indeed been found to respond rapidly to learning episodes (Xu, Yu et al. 2009).

<sup>177</sup> Cognitive flexibility as measured by drawing non-existent objects was shown to be greater in bilingual children, suggesting an interrepresentational flexibility that relies on representational redescription processes fostered by bilingualism (Adi-Japha, Berberich-Artzi et al. 2010).

that make it attentive and extra sensitive to specific environmental inputs, but it is the content of these inputs that partly determine further brain development and cognitive processes (Karmiloff-Smith 1992). Indeed, the importance of such interactions between environment and cognitive processes applies to most developing cognitive or behavioral functions, ranging from motor performance to drawing, reasoning about physics and symbol manipulation (Elman, Bates et al. 1997). Imprinting in chicks being considered a very simple example of such interaction (Elman, Bates et al. 1997), it is of limited relevance for our purposes compared to forms of learning that involve explicit instruction.<sup>178</sup>

The mastery of speech and symbol manipulation has an important impact on learning and associated processes. Research in bilingual children, for example, confirms that the mastery of more than just a single language provides these children with increased flexibility in cognition and behavior, probably depending on their being able to shift between two different representations of a single object or task (Adi-Japha, Berberich-Artzi et al. 2010 ; Bialystok, Shenfield et al. 2000). More than just generally facilitating learning and development via environmental inputs, these inputs provide subjects with extra resources to influence their other cognitive and behavioral processes.<sup>179</sup> For if development and learning rely in a crucial way on a series of Representational Redescriptions, then it should be possible to manipulate or reconfigure the representations involved such that they have specific influences on the learning processes and later outcomes.<sup>180</sup> Indeed, the acquired capability of explicitly manipulating and redescribing one's representations for a particular task yields many benefits for quick learning, correcting, adjusting, and cross-domain transfer of contents – and environmental information can play an important role in this capability (Clark and Karmiloff-Smith 1993).<sup>181</sup> Concurring with this, Hollis & Low conclude

---

<sup>178</sup> As noted earlier, imprinting in chicks has been interpreted as a demonstration that the strict distinction between innate and acquired constraints on cognitive processes is not useful. Though Wimsatt approached the matter differently from the modularization account, his account of the generative entrenchment of environmental information in a complex and modular system does correspond to a large extent with the former (Wimsatt 1986).

<sup>179</sup> However, such utility is only available at the developmental or learning stage where the relevant information has obtained the necessary representational status and not earlier (Karmiloff-Smith 1992).

<sup>180</sup> The authors of 'Neuroconstructivism' even describe a feed-back loop between a proactive child that affects its environment on the basis of its representations, eventually initiating further interactions that subsequently affect its own environmental inputs and thus its cognition and development (Sirois, Spratling et al. 2008).

<sup>181</sup> Language, symbols and narratives, for example, are being used for such manipulations and redescriptions. This will be discussed more specifically in chapters II.4 and III.4.

after their discussion of the effects of instruction and visual examples on children's drawings that: "[a] hierarchically organized exemplar training program may assist the progress of representational redescription by decomposing the task into an ordered series of subfeatures that partition the problem-solving space" (Hollis and Low 2005 641). Hence, external information not only speeds up learning but also influences the decomposition of the task's representation. As a result, the kludge formation that accounts for the behavioral mastery and flexibility pertaining to the task, involves the integration of environmental information. We will consider in Part III further examples of how such targeted analysis and manipulation of action representations by an agent can amount to his putting constraints on the space of actions available to him. Indeed, we will argue that expertise with particular intentional actions can imply kludge formation in underlying mechanisms, further sculpting the agent's space of actions – as when an opera singer's preferences in gender relations tacitly transpire in his performance of Don Giovanni.

### **2.3 Modularization considered as a process of kludge formation**

After our comparison of the neuroconstructivist account of modularization with our seven characteristics of kludge formation, it seems not inappropriate to consider the modularization process as a form of kludge formation. Modularization is a process of increased encapsulation of a domain's – or subdomain's - information, being processed with an ever diminishing influence of other brain processes. The process is mainly observable via changes in cognitive and behavioral responses, with variability between stages and between subjects potentially occurring during this developmental process. This variability partly comes about through environmental information – like verbal instruction or examples – that plays a role during modularization and that can affect components of the task or the task as a whole. For such environmental information to be effective, a representational redescription of the relevant knowledge must already have obtained before in order to enable the subject to adjust its responses: the subject must already possess a degree of behavioral mastery before he can flexibly adjust to the new information. It is this complex interaction between the initial proceduralization of knowledge and its subsequent explicitation that is characteristic for this account and which offers some suggestions to our further development of the notion of a kludge.

First and most importantly, according to this account human subjects usually develop more than just a single representation of a particular domain of knowledge or behavior. These representations have different formats and accordingly also different properties. Some are more suitable for automatized motor performance while others, when redescription in a more abstract format has taken place, lend themselves better

for explicitation and flexible adaptation. The simultaneous availability of a plurality of representations pertaining to a particular task can explain surprising cognitive or behavioral results. This is the case when an agent shows temporary behavioral regress during a stage transition or when he can perform one and the same task according to different modes, each with different properties. In Part III, we will consider whether the same situation holds for the action determination processes on which we will be focusing our attention.

Second, given the presence of multiple representations after development and learning a task, it remains to be considered whether kludge formation can only obtain for the early stage of learning, or whether the more explicit representations that are developed at later stages of learning can also correspond with kludge formation. After reaching the stage where an agent can explicitly correct and adjust his task performance, may the adjusted representation in turn become proceduralized as well? Since neuroconstructivist research of modularization focuses mainly on children and development, the present account does not offer a clear position in this regard, although Karmiloff-Smith's account of herself learning to play the piano or solve Rubick's cube confirms that subsequent proceduralization can still occur in adults (Karmiloff-Smith 1992). This could theoretically lead to a situation where different implicit representations exist parallel to each other: one that emerged during the child's early learning and a second one that developed after explicitation, resulted again in an adjusted task representation.

With these two final remarks, we are already embarking on a discussion of so-called 'dual-process theories', which focus on the presence of two distinct types of processes underlying many cognitive and behavioral functions. Such processes are distinguishable in many respects, though they share largely overlapping domains of activation. Thus, as dual-process research demonstrates, it can occur that the two processes are provoked by the same environmental stimulus, thereby yielding two distinct behavioral responses to a single stimulus. This has led some authors to point out that humans have in fact 'two minds in one brain' (Evans 2003) which perform according to different types of processes, one of which is even labeled a 'cognitive monster' (Bargh 1999) as it performs its task more or less automatically. Particularly the latter phrase gives air to the negative assessment of this automatized type of processing, an assessment that is not common in the context of the proceduralization research, even though there are some similarities between the automatized and the proceduralized types of task performance. In the next chapters we will further explore the dual-process theories and consider whether they, too, in fact concern cases of kludge formation and thus bear witness to the capability of developing more than

one mode of performing a task once it is practiced regularly. If so, these theories may add some insight in the process of sculpting the space of actions as it happens with increasing expertise.