



UvA-DARE (Digital Academic Repository)

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

Keestra, M.

[Link to publication](#)

Citation for published version (APA):

Keestra, M. (2014). Sculpting the space of actions: explaining human action by integrating intentions and mechanisms. 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, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Contents

Introduction: intentional action and a sculpted space of actions	2
How aristotle avoids the paradox of the expert by accepting causal pluralism	5
The importance of development and learning in aristotle's account	6
Conceptual innovation in aristotle's account of the development of expertise	9
Part I: investigating four methods of explanation in cognitive neuroscience	11
Part II: dynamic cognitive mechanisms and their stable adjustments due to development and learning	13
Part III: sculpting the space of actions for the performance of intentional expert action	15
Part I – Conceiving and explaining: an intricate relation	21
1 Introduction: a common capability with divergent results	22
2 Concepts as delineations for empirical content	29
2.1 Concepts as 'clean instruments' for neuroscience	30
2.2 Connective analysis and ascription criteria	33
2.3 Non-convergent and variable criteria, and their implications	37
2.4 Heuristic use of conceptual divergences, yet with limitations	42
3 David Marr and the involvement of concepts in multi-level explanations	47
3.1 The analysis of computations or tasks – not concepts - should guide scientific investigations	48
3.2 Constraints that co-determine the computations' appropriateness	52
3.3 Two further levels for multi-level explanations	55
3.3.1 <i>The algorithmic level and the representation of information</i>	55
3.3.2 <i>The implementation level and neuroscientific evidence</i>	58
3.4 A loose interdependency between levels	61
3.5 Modularity and some limitations of Marr's methodology	63
4 Modest, all too modest: the search for neural correlates	68
4.1 Identifying a minimal yet sufficient neural correlate – of what?	69
4.2 Further limiting the phenomenon and its correlates	74
4.3 Lessons on explanation drawn from the NCC research	77
5 Mechanistic explanation and the integration of insights	81
5.1 From the mechanistic world picture to the method of mechanistic explanation	85
5.2 Memory and the mechanistic explanation of learning	89
5.3 Defining the phenomenon as a first step	94
5.4 Facilitating the explanatory task by decomposing the phenomenon	96
5.5 Localization of the decomposed phenomenon	100
5.6 Mechanistic explanation and mechanism modifying dynamics	105
5.7 Some limitations of the mechanistic explanatory approach	110
6 Concluding remarks after considering the four methodologies Dynamics of change and stability in cognitive mechanisms	113

Part II – Dynamics of change and stability in cognitive mechanisms	117
1 Introduction: from dynamics to stability and back again	118
1.1 Kludges: mechanism adjustments and expansions	121
2 Modularization as a process corresponding to learning and cognitive development	126
2.1 Neuroconstructivism and the relevance of modularization	128
2.2 Modularization and the seven kludge characteristics	130
2.3 Modularization considered as a process of kludge formation	138
3 Dual-process theories and a competition between forms of processing	141
3.1 Distinguishing between forms of processing, irrespective of tasks?	144
3.1.1 <i>Considerations of the distinction between automatic and controlled processes</i>	148
3.1.2 <i>Processing limitations held responsible for the distinction between automatic and controlled processing</i>	150
3.1.3 <i>Memory systems invoked for the explanation of the distinction between automatic and controlled processing</i>	152
3.1.4 <i>Some strategies that allow a shift between automatic and controlled processing</i>	153
3.1.5 <i>Representational differences and the shift between automatic and controlled processing</i>	155
3.2 Automatization of controlled self-regulation and the seven kludge characteristics	161
4 The brain as a mechanism capable of kludge formation and open to external information	166
4.1 Symbols, simulators and the malleability and stability of cognitive processing	168
4.2 Simulators and the kludge characteristics	174
4.3 Reaching outside the skull: how can external objects become integrated?	179
4.4 Cognition-extensions and the kludge characteristics	188
5 Dynamic mental mechanisms, kludge formation and establishing constraints on the space of options	198
Part III – Sculpting the Space of Actions with Intentions and Mechanisms	203
1 Introduction: multiple mechanisms yet stable patterns	204
1.1 ‘Sculpting the space of actions’ – an important ingredient for the explanation of expert action	207
1.2 Determining an action via a cascade of intentions	216
1.3 The cascade of intentions and a sculpted space of actions	222
2 Motor intentions: the first step in the hierarchy, or not?	226
2.1 A philosophical analysis of motor intentions and guidance	228
2.2 Motor intentions and chunks: evidence about developing complexity at the bottom of the hierarchy	234

2.2.1	<i>Motor intentions, representations and the role of expertise</i>	237
2.2.2	<i>Motor intentions and mechanisms that change with growing expertise</i>	243
2.2.3	<i>Motor intentions and differential generative entrenchment of components</i>	247
2.2.4	<i>Motor intentions and consistency of action</i>	250
2.2.5	<i>Motor intentions and some evidence concerning their neural implementation</i>	253
3	Proximal intentions: a mediating role	257
3.1	A philosophical analysis of proximal intentions	259
3.1.1	<i>Resolving conflicts between action options</i>	260
3.1.2	<i>Proximal intentions and blocking habitual action</i>	263
3.1.3	<i>Proximal intentions and constraints for anchoring an action</i>	265
3.1.4	<i>Proximal intentions and their peculiar position in the agent's psychology</i>	267
3.2	Proximal intentions and cognitive mechanisms that determine anchored actions	270
3.2.1	<i>Processes for the resolution of conflicts between action options</i>	276
3.2.2	<i>Blocking a habitual action according to the CS/SAS model</i>	280
3.2.3	<i>Multiple processes and the constraints for anchoring an intention in a situation</i>	284
3.2.4	<i>Proximal intentions and some evidence concerning their neural implementation</i>	287
4	Distal intentions: governing the intentional cascade?	297
4.1	A philosophical analysis of distal intentions	303
4.1.1	<i>Closing the gap between distal and proximal intentions and sculpting the space of actions</i>	304
4.1.2	<i>Hierarchy and stability of a planning agent's web of intentions</i>	307
4.1.3	<i>A role for the imagination in weaving one's comprehensive web of intentions?</i>	309
4.1.4	<i>Narrative simulation as an additional resource of establishing one's agency</i>	312
4.1.5	<i>Narrative as simulation of distal intention, contributing to a sculpted space of actions</i>	318
4.2	Distal intentions and narratives and their reciprocal interactions with cognitive mechanisms	321
4.2.1	<i>From memorized experiences to the simulation of future actions</i>	324
4.2.2	<i>Distal intentions in the intentional cascade: from action control to mental time travel</i>	328
4.2.3	<i>Simulation and the flexible imagination of a future action</i>	332
4.2.4	<i>Narrative and additional benefits of the simulation of multiple distal intentions</i>	336
4.2.5	<i>The socio-cultural nature of some schemas for narrative simulation</i>	343
4.2.6	<i>Narrative simulation and some evidence for its implementation</i>	349
	Conclusion and Summary. Why sculpting the Space of actions matters	356
	Figures: I, II, III	371
	References	377
	Samenvatting	432

Contents

Part I – Conceiving and explaining: an intricate relation

21

Being the first of three parts, Part I concerns methodology. Explaining someone's cognitive and behavioral performance requires showing how the brain makes it possible, while internal and external sources of information play an important role. Such an explanation can only adequately be developed by first defining the phenomenon that is to be explained. Consequently, an explanatory method is proposed that facilitates an interdisciplinary investigation, integrating insights from various empirical sciences and philosophical analysis.

1 Introduction: a common capability with divergent results

22

Human agents are capable of learning a wide range of actions, some of which require a lot of expertise, like performing an opera role, while other actions can be carried out impromptu. Besides, some actions require explicit attention and conscious coordination, while others are realized more automatically. This book explains how we can understand and explain the fact that an expert's automatic actions, too, can still be considered intentional and subject to the coordination and organization of his actions. From Aristotle onward, philosophers and scientists have had an interest in explaining how an individual agent's behavioral repertoire, or "space of actions", is sculpted via a plurality of processes and how this is visible in his actions. 'Sculpting a space of actions' implies increasing the differentiation between habituated versus unpractised actions, between preferred versus avoided actions, further increasing the consistence and coherence between his actions. Explicit instruction by teachers, individual deliberate practice, endless repetition of motor actions and more all contribute to this sculpting process which affects behavioral, cognitive and neural processes. Given the number and diversity of these determining factors and the additional complexity of their interactions, it is evident that a complex explanation is required. Part I scrutinizes four different explanatory models, looking for a model that can cope with the diversity of determining factors and also account for the dynamics involved in the process of an agent's sculpting the space of his actions.

2 Concepts as delineations for empirical content

29

Explaining a complex and dynamic phenomenon like the process of an agent's sculpting the space of his actions requires some integration of conceptual insights with factual, empirical evidence. According to the model about to be discussed in this chapter a strict and logical distinction between concepts and facts should be recognized, in which concepts require a philosophical analysis whereas facts are the results of scientific investigation. Importantly, this account contends that it is possible

to provide a consistent and clear conceptual framework of psychological functions based upon the analysis of concept use and the behavioral criteria that are commonly used to ascribe someone a particular function. Empirical scientists who do not understand nor comply to such a framework are accordingly liable to utter nonsense when presenting their findings as evidence concerning psychological functions. However, our critical discussion points out several weaknesses in the assumptions behind this account and suggests instead that conceptual ambiguities and divergences can be exploited as heuristics for empirical investigations and that a more pluralist approach to the relation between concepts and facts should be allowed.

3 David Marr and the involvement of concepts in multi-level explanations 47

The influential approach to cognitive neuroscience developed by Marr explicitly acknowledges that explanations are the result of a plurality of insights which are different in kind. The account involves three different levels of analysis or levels of explanation called computational, algorithmic and (neural) implementation levels. These offer different explanatory perspectives on a particular cognitive task, which are only loosely interdependent according to Marr. The computational level concerns the goal and functionality of the task, including the logic of it in light of its wider context. The algorithmic level is devoted to options for the representation of information used in the task and its transformation during task performances. Investigation of the implementation of the task, like in the brain or in a computer, can subsequently help to constrain the options for the algorithmic level and vice versa: some algorithms would be better served with a particular implementation than others. Although this account is explicit in its acceptance of explanatory pluralism, it leaves undetermined how the different explanatory perspectives on a cognitive function can be integrated and result in a more comprehensive explanation. Moreover, the approach allegedly has difficulties with complex and interactive systems which are subject to multiple influences - such as the kind of systems that allow agents to become experts in singing and moral actions.

4 Modest, all too modest: the search for neural correlates 68

Another explanatory model is employed in consciousness research. Since consciousness is a complex phenomenon and lacking agreement about a conceptual framework for it, research of consciousness focuses largely on two types of research. On the one hand it focuses on the investigation of small transitions from a particular content into consciousness or out of it, while on the other hand it focuses on the investigation of different background states of consciousness like coma, sleep and conscious states.

Scientists then look for specific neural activations that can be correlated to such transitions or states: neural correlates of consciousness. Modestly avoiding a strict conceptual definition or task description of consciousness, they hope that from assembling many of its neural correlates a general account of consciousness emerges - possibly a functional explanation that might offer a substitute for a definition, or an overlapping neural process. However, two problems remain: firstly, without any preliminary definition of consciousness it is impossible to accept or reject a finding as being a neural correlate of consciousness and secondly, upon accepting a neural correlate we still don't possess an explanation of how it contributes to consciousness. For that we need to specify the explanatory mechanism.

5 Mechanistic explanation and the integration of insights

81

A model that recognizes differences between partial explanatory accounts yet provides resources for their integration, is mechanistic explanation - which is not the same as classical mechanistic thought. Mechanistic explanation explicitly requires as a first step the preliminary definition or delineation of a phenomenon, such as for example a cognitive task. Secondly, the cognitive task must be - if only tentatively - decomposed into component tasks. In some cases, a component task can further be subdivided into even smaller tasks, as has successfully been done with vision or memory. A third and final heuristic that researchers carry out is the localization of the respective tasks somewhere in the responsible organism or system. Applying these three heuristics, researchers can uncover an explanatory mechanism that is responsible for the task, consisting of component parts and operations at different mechanism levels. These components interact in an organized fashion and in response to both internal and external conditions. Importantly, changes regarding a cognitive task that occur during development and learning depend upon changes that affect such an explanatory mechanism. Such changes can correspond either with the recruitment of a particular new mechanism components, or with a novel configuration of components, or with the emergence of new types of interaction with the environment, or a combination of such mechanism modifications.

6 Concluding remarks after considering the four methodologies

113

Part II - Dynamics of change and stability in cognitive mechanisms

117

Part II is concerned with the explanation of structural changes in an agent's behavioral and cognitive performance. Insights derived from the method of 'mechanistic

explanation' from Part I are applied to four different theories about development and learning. With these results we can go on to explain how a 'sculpting process' can have enduring effects on the mechanism responsible for changes in expert performance.

1 Introduction: from dynamics to stability and back again 118

The previous Part has ended with an analysis of how development and learning corresponds with the modification of a relevant explanatory mechanism. Part II subsequently discusses different prominent explanations for changes in an individual's cognition and behavior as a result of such modifications. It is argued that such changes further sculpt an agent's space of action, for example by expanding the agent's ability for different modes of action, like when he develops both an automatic and consciously controlled mode of performing a particular action. Such an expansion is part of what distinguishes an expert, who is better able to exert control on his modes of action, from a novice. More specifically, it is argued that in many cases we can observe the formation of a 'kludge', or an extra component that is 'cobbled-together', in a mechanism, which can explain the stability of effects of such changes. Seven characteristics of such kludges are discussed, to which we will return when we will consider different explanations of the consequences of learning and development in terms of kludge formation, affecting an explanatory mechanism.

2 Modularization as a process corresponding to learning and cognitive development 126

According to neuroconstructivist accounts, development and learning correspond with increasing modularization of underlying neural processes. This modularization can be observed in both the proceduralization of a skill - it becoming more stable and flexible while also increasingly automatized and implicit - and its subsequent explicitation - rendering the skill eventually accessible for explicit correction or transformation. These developments are not just constituted by changes in the neural processes but also by the changes in the representations involved as Representational Redescription occurs. This neuroconstructivist theory of modularization can be understood largely in terms of kludge formation, affecting the explanatory mechanism underlying a skill in several ways. In other words, kludge formation can be observed in the agent's actions and involves changes in the representation of information involved as well as modification of the underlying mechanism, confirming the methodological insights presented in Part I.

3 Dual-process theories and a competition between forms of processing 141

Development and learning produce differences in an agent's performance and its underlying mechanism. Yet in many cases, the result is that an agent has more than just a single mode of performing an action at his disposal. Prominent in cognitive sciences, dual-process theories contend that an agent can perform many cognitive and behavioral tasks not just by a single type of processing but by two different types: automatic and controlled processing. A task's underlying process can become automatized after some time, which affects important properties of the task performance as it no longer requires the involvement of conscious control, nor explicit representation of the task and is usually faster. This shift from controlled to automatic processing can be partly described in terms of kludge formation in which the underlying mechanism is modified, associated with changes in the representations involved. An agent might employ several available strategies for controlling the type of processing by shifting his attention, by preliminary activating a behavioral schema, or by other ways. Such controlled self-regulation can itself become automatized, which can be partly understood in terms of kludge formation as well. Kludge formation thus sculpts the agent's space of actions, contributing to the varieties in his performances.

4 The brain as a mechanism capable of kludge formation and open to external information 166

We've learnt that an agent's performance can rely on several modes of processing, to which development and learning contribute as these result in kludge formation. In this third account of learning and development we will apply this notion to the emergence of a 'simulator', which essentially consist of a complex representation that is composed of components and is not stored as a whole but in a distributed manner across the brain. Such a simulator can be involved in various cognitive processes, facilitating an agent's perception, imagination and action concerning that representation, for example. Such a representation allows for recomposition or redescription and for the inclusion of component representations of environmental objects, like tools, or information, like language. Due to this, an object or word can also activate or compose a simulator in an agent, which in turn produces the 're-enactment' of previous states of perception, motor activity or cognition, affecting in many ways his subsequent performance. The notion of a simulator largely matches with our notion of kludge, like it having effects on task performance, its being composable from previously established components, its integrating environmental information. In addition, establishing such simulators further contributes to the agent's sculpted space of actions.

5 Dynamic mental mechanisms, kludge formation and establishing constraints on the space of options 198

While focusing on multiple processes like child development and skill acquisition, the shift from conscious to automatic control, learning to use tools and language and the like, Part II has confirmed how mechanisms responsible for particular functions can become modified during those processes. Such mechanism modification often implies kludge formation which has been shown to involve multiple characteristics. This process has consequences for the agent's performances as it contributes to his 'sculpted space of actions': this space can become to include novel actions while excluding others, some action options will become easily activated while others do not, the relations between actions and particular environmental conditions might change, and so on. As a result of this sculpted space, an agent's action performance can acquire a certain stability and consistence, even if many of his actions are performed through another type of processing than consciously controlled action. In addition, as an agent's expertise to a large extent depends on his having established a sculpted space of actions, this space is also involved in his capability to adjust his actions in ways that a novice can not do.

Part III Sculpting the Space of Actions with Intentions and Mechanisms 203

Part III is devoted to the explanation of expert action, such action being in many senses different from novice action. Differences between experts and novices are explained in terms of their action intentions, which are elaborated far more, but also in terms of underlying cognitive and brain processes. It is explained how experts, enabled by the process of 'sculpting the space of actions', can perform increasingly complex actions while coordinating and organizing these much better than novices.

1 Introduction: multiple mechanisms yet stable patterns 204

The same action can be produced by multiple non-identical mechanisms and mechanistic explanation helps us to account for the changes in an agent's action performance and control that occur as he gains expertise. An explanatory mechanism can become modified when kludges emerge, often by a reconfiguration of previously present mechanism components and while somehow integrating environmental objects or information. In section III.1.1. we clarify what 'sculpting the space of actions' entails and how it contributes to an agent's complex but relatively stable 'sculpted space of actions' and its internal structure. These insights can help to analyze and explain from both a philosophical and a cognitive neuroscientific perspective how different types

of intention are involved in the agent's actions and in his action planning. Indeed, we will consider the presence of a hierarchical 'intentional cascade', consisting of motor, proximal and distal intentions with each type having specific properties, involving specific representations and neural implementations. The three types of intentions play a particular role in an agent's intentional actions and also interact in several ways. Scrutinizing this intentional cascade will learn us a lot about the sculpted space of actions that an agent has.

2 Motor intentions: the first step in the hierarchy, or not?

226

Motor intentions are held responsible for the implicit guidance and adjustment of intentional motor actions, which are distinct from mere motor reflexes. Other than reflexes motor intentions rely on action representations that contain information about motor movements and relevant environmental conditions, which agents are capable of storing in memory. These representations are in non-conceptual form yet they are structured and are modifiable. Expertise consists in part in learning to compress or chunk motor representations and to gather many of those chunked motor representations, thus sculpting a space of actions and enhancing consistency between actions. Like simulators, these representations influence multiple cognitive processes, enabling an expert to recognize and respond faster and more flexible to relevant environmental information than a novice can. The mechanism modification at neural levels associated with expertise is twofold: at first the neural processes become more efficient, then they can become associated with other processes, allowing an expert more complexity in his actions and also leaving room for taking his other intentions into account.

3 Proximal intentions: a mediating role

257

Proximal intentions mediate between the distal intentions that contain representations of future actions in a conceptual format and the motor intentions that guide motor movements in response to environmental conditions. A proximal intention is responsible for quickly anchoring or perhaps instead inhibiting a distal intention in a concrete perceived situation, specifying the necessary representation components, partly based upon an agent's stored motor representations. An expert can rely upon his having assembled many relevant and complex representations, or action schemas, enabling him to usually act more quickly and adequately than a novice but also to exert more control over his actions. A mechanistic explanation of these effects of acquired expertise involves a dual process theory, consisting of an automatic 'contention scheduling' process that can be to some extent modulated by controlled supervisory

processes, each relying on distinct neural mechanism components that can interact. Expertise then implies the agent's familiarity with these processes, providing him with several options for determining and constraining his proximal intentions in multiple ways that allow him to let these fulfill their mediating role optimally.

4 Distal intentions: governing the intentional cascade?

297

On top of the hierarchical 'intentional cascade' are distal intentions that are in a propositional format and more abstract, requiring further anchoring and specification for their future execution by lower levels of intention. They are held responsible for governing and coordinating an agent's actions and help him to foster consistency in his actions by taking his wider web of intentions and his future actions into account instead of focusing just on a single intention in a specific situation. In order to be effective, distal intentions must be able to multifariously interact with his proximal and motor intentions and influence his sculpted space of actions, too. Agents typically do so by a narrative simulation of future situations or a more comprehensive narrative self-account. Such a narrative simulation consists of complex representations of action at several levels of hierarchy, employing the previously mentioned simulators that are stored in a distributed way across the brain and reconfiguring these in novel ways. An expert has learnt how to determine and employ his distal intentions more effectively than a novice, also by including specific schemas and cultural ingredients. Distal intentions exert their influence on future actions partly via modulation of the agent's neural 'default mode network' which has rich connections to neural networks involved in cognitive, affective and mentalizing tasks that are relevant for determining his future actions. Though not without limitations, distal intentions can thus influence these actions in several ways as they become entrenched in that network. This completes the rich interactions between the agent's intentional cascade and his sculpted space of actions and enables him to indeed become an expert - an expert who causes so often surprises and invites complex interdisciplinary explanation.

Conclusion and summary. Why sculpting the space of actions matters	356
Figures: I, II, III	371
References	377
Samenvatting	432