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

Keestra, M.

Publication date
2014

3 PROXIMAL INTENTIONS: A MEDIATING ROLE

Above, in chapter III.2, we pointed out how amazingly controlled and intentional an expert singer performs his Don Giovanni, even though the complexity and comprehensiveness of his acting, singing, observing and interacting is such that we cannot expect him to deliberate and rationally decide about all of it on the spot. Instead, so we argued, his behavior is facilitated by a sculpted space of actions partly determined by implicit motor intentions, knowledge structures consisting of representations of complex patterns and associated motor movements, assembled over time with practicing, learning and experience. However, if only motor intentions would determine this sculpted space of action, this would raise several questions.

First, one could ask whether a sculpted space, determined by motor intentions alone, would facilitate the agent's task of selecting a single action option from the many motor intentions that a situation with many affordances may activate. Are there constraints available that help him to choose between those options, taking into account that motor intentions operate at a temporal scale that makes it impossible for deliberate and conscious choice to intervene? Constraints that are intimately related to his previous choices and experiences and can therefore be considered to be in line with these?[365]

Second, even if there are no alternative action options available, will a motor intention automatically ensue into an action given certain affordances, similar to those automatized actions that escape any form of control? Although Aristotle already argued in favor of habits as part of an agent's moral behavior, there are situations imaginable in which these habits are morally inappropriate. Therefore the question rises whether it is possible to block a motor intention's application under certain circumstances?

Third, apart from blocking a motor intention in an exceptional situation, are there perhaps further constraints on an agent's sculpted space of actions? For the absence of further constraints on motor intentions could lead an agent to act inconsistently over time or in ways that do not cohere with his other intentions or beliefs. Indeed, can we consider those other intentions and beliefs to contribute to an agent's personality in such a way that they further constrain his motor intentions as these on their own probably cannot constrain each other adequately?

For example, consider our expert singer who has to avoid confusion in an on-

[365] The role of the agent's history has become – again – important in the philosophical debate about his autonomy of action. With such history being given its due, autonomy becomes a more complex issue as former choices and experiences may contribute to ongoing and future actions even if such contributions are not always consciously deliberated.

stage situation with many stage props and persons that could offer various action options, pertaining to both his Saint François and Don Giovanni roles. How does he decide to which affordances he should respond and which intention - out of many intentions - to act he should implement in a given situation? Moreover, consider the situation of an opera star who is flown in for just one guest appearance in a regular series of performances of Don Giovanni. Being familiar with the score and libretto, he must still find himself at home in a new direction and stage, and so on. Under an exceptional director like Peter Sellars, who invited as singers for Don Giovanni and his servant Leporello twin brothers and set them as drugs dealers in Harlem, he may have to block some habitual actions while substituting them for newly learnt actions. Of course, he must prepare his role by looking at pictures of the set design, reading stage directions and formulating for himself intentions to act in certain ways or to avoid certain actions – depending upon his expertise with the role and with other directions. Having prepared himself in this way, the experienced singer may need only a short rehearsal of some crucial scenes to practice the performance satisfactorily. This rehearsal would allow him to put additional constraints on the set of motor intentions that his preparations alone could probably not do.

Clearly, distal intentions are involved in the singer's preparation by looking and reading and planning, yet proximal intentions must also be established, for example by rehearsing. These additional intentions further constrain and sculpt the agent's space of actions. With motor intentions being processed in a semi-modular way and being non-conceptually and unconsciously responsible for the transformation of specific perceptual information to corresponding sensory-motor information in a given situation, distal and proximal intentions offer further constraints. With regard to proximal intentions, Pacherie writes that, the: "problem at the level of P-intentions consists in integrating conceptual information about intended action inherited from the D-intention with perceptual information about the current situation and memory information about one's motor repertoire to yield a more definite representation of the action to be performed" (Pacherie 2008 185). Integration of these sources of information leads to the 'anchoring' of the distal intention in a given situation, which is an important task of proximal intentions.

Is this task of specifying the more abstract distal intentions for their eventual initiation and performance alone a crucial task, proximal intentions are allegedly also involved in monitoring and guiding the outcome of the action. These tasks imply the interaction between intentional contents that are partly hierarchically related to each other, and processes that may be differently structured, being constrained by those contents and by environmental and motor conditions. Execution of these tasks requires processes that probably are quite different yet still have to be related to each other. Next to the processes corresponding to motor intentions, proximal intentions contribute to the determination of action as well. Since the time scale of proximal intentions is not so restricted as is the case in motor intentions, explicit and conscious perception and cognition are able to play a role, according to this framework. As a result, rational constraints in the sense of coherence and consistency constraints are at work at this level, too, with proximal intentions being responsible for keeping track of an optimal action performance and for controlling for potential side effects (Pacherie 2006 ; Pacherie 2008) – things that motor intentions are not capable of.

In as much as proximal intentions play an important role with regard to all the issues raised above, they contribute in important ways to the sculpted space we're investigating. As noted before, they don't do so in separation, as they meanwhile integrate also contents from higher and lower level intentions. So now we will first provide a philosophical analysis of this complex task and an argument for its importance, after which we will be turning to empirical evidence.

3.1 A philosophical analysis of proximal intentions

We started the sections on motor intentions above with a philosophical analysis of motor intentions, largely leaning on Frankfurt's account of guidance. Now that we're shifting our focus on proximal intentions and their intermediate position, it will be Bratman's conception of present-directed intentions that guides our analysis, partly because Pacherie has explicitly built her framework on this conception.[366] However, it has been Frankfurt's introduction of the distinction between different orders or levels of intentions into his account of action that has influenced both (Frankfurt 1971). Frankfurt has argued for these levels of intentions and their interaction as a structure that is required to regulate an agent's actions. Moreover, he argues, it is this structured interaction between an agent's intentions that reflect his identity as a person (Frankfurt 1988). Bratman has taken up important elements from this account of intentionality by way of a structured interaction but has elaborated upon its dynamics and has developed a philosophical analysis that will turn out to be quite suitable for integration in an explanatory account of action. This is due to a large extent to his greater emphasis on temporal dynamics of the structured interaction (Bratman 1987).

[366] Proximal and distal intentions are terms that are being used by Mele, with proximal intentions referring to intentions for the 'specious present'. Bratman introduced the notions of future-directed and present-directed intentions for comparable purposes as proximal and distal intentions. Mele relativizes differences between the positions and simply notes in that context that 'terminology varies' (Mele and Moser 1994 65, footnote 10).

With that addition, the philosophical analysis has more to offer to an account like ours, which aims to contribute to a mechanistic explanation of action determination. Given that for our account we have a special interest in the dynamics involved and in the possibility of an emerging sculpted space of actions, alleviating in turn some of the tasks involved in determining the agent's appropriate action, Bratman's work will be more closely discussed in what follows.

3.1.1 Resolving conflicts between action options

In section III.2.1., we found Frankfurt analyzing guidance in terms of ongoing action monitoring and intervention mechanisms that remain passive as long as their intervention is unnecessary (Frankfurt 1988 75). In the current context, we find him interested in the different degrees of commitment or engagement an agent can enjoy regarding his decision to perform any action at all. In both cases, gradual variety obtains for an internal process that is relevant for an agent's actions.

Appreciating the fact that in any given situation an agent may have the troublesome experience of having multiple desires to perform certain actions, Frankfurt has introduced the notion of different orders of desire or volition. An agent may in such a case not be content with the desire that has eventually won out the competition, in which case: "he wants to be motivated effectively, with respect to the alternatives he faces, by some desire other than the one that actually moves him to act as he does" (Frankfurt 1988 48). The agent's internal conflict can therefore be twofold.[367] First, the conflict may be among his multiple first-order desires, each of which concerns a particular option to act.

Second, another internal conflict would occur when a first-order desire has eventually won out the competition and would determine the agent's action but turns out to be in conflict with a second-order volition that he also has. In that case the latter volition would have preferred the action to be determined by an alternative desire. The question is what role these conflicts and their solutions play in a given situation where several action options are at stake and where an agent can only execute a single action. Are there strategies or constraints available that facilitate solving these recurrent problems? Is an agent always forced to engage in a sequence of choices, in which he endures and solves the first type of conflict and then subsequently deals with the second type? Or are there other forms of interaction between the different orders

[367] An important feature of Frankfurt's account is that it is not uncommon for an agent that: "no second-order volition plays a role in the economy of his desires" (Frankfurt 1988 50). The phrase 'economy of desires' alludes to Frankfurt's view that in many cases, rational deliberation is not really at stake, but rather a competition between first-order desires in terms of the costs and benefits involved in their realization.

possible that would alleviate these tasks in some way or another?[368]

In a situation where different desires to act cannot be realized simultaneously, we are usually required to reject all except one in order to resolve this conflict. A simple solution to reach this result would be by ordering the desires and establishing a preferential ranking of action options, being members of the same order. These options then differ in degree rather than in kind for the agent. However, in a given situation it may be quite a large cognitive task to articulate and then order all action options, provided they would all be comparable to each other. Imposing a constraint that divides these options into different categories could be helpful in constraining this task. This is the case when an agent has specified a set of second-order desires and then identifies only with those action options that conform to these. Conversely, he may have decided that a particular desire should not belong to his particular space of action options and: "is finally to be excluded from the order of candidates for satisfaction" (Frankfurt 1988 68). The latter case refers to a situation in which the agent has put the rejected desires external to himself as a person and in which it would be profitable for him to sculpt his space of actions correspondingly.[369]

This analysis does not comprehensively show how an agent goes about to select and realize – anchor - one of his desires or intentions in a given situation but it does contribute to answering that question. Although Frankfurt focuses especially on the issue of an agent's responsibility for his actions, his analysis does also imply an interest in the contribution of the enduring structure of the agent's personality and identity to solving this question.[370] In his analysis he aims to demonstrate how this structure does constrain the potentially large cognitive task of selecting these action options among competitors and then realizing them. By deciding about the order of his desires and

---

[368] This does not mean that second-order volitions do not influence the competition between first-order intentions. Indeed, when Frankfurt later analyzes the wholeheartedness with which an agent can embrace his own beliefs or attitudes or intentions, he contends that the agent's satisfaction is with: "these psychic elements (…) rather than others that inherently (i.e. non-contingently) conflict with them, should be among the causes and considerations that determine his cognitive, affective, attitudinal, and behavioral processes" (Frankfurt 1999 103). This satisfaction implies that the agent will not resist or reconsider the outcome of the first type of conflict, between first-order intentions.

[369] It has been argued critically that Frankfurt eschews to assign a central role to objective or rational criteria for such decisions, which raises doubts about both the role of rationality in such an agent and about the moral value of them (Buss 2002). Frankfurt does indeed doubt about the nature of these decisions, as he explicitly admits (Frankfurt 1988 68). Moreover, in his reply to Buss he contends that it is well conceivable that living an immoral life may be good from the person's perspective, even if it is Hitler – so his decisions may be rational yet immoral, indeed (Frankfurt 2002b).

[370] Although as far as we are aware of, they do nowhere explicitly discuss each other's work, there is some affinity between this investment in the person as a final cornerstone of analytical philosophy of action and Ricoeur's phenomenological and hermeneutical philosophy of action. Indeed, the latter explicitly acknowledges that "the *question of personal identity* is posed at the point of intersection between the two philosophical traditions" (Ricoeur 1992 17, italics in original).

particularly by including and excluding some of these intentions from the space of actions that he is satisfied with, the agent constrains and guides these tasks: "[o]ne thing a deliberate decision accomplishes, when it creates an intention, is to establish a constraint by which other preferences and decisions are to be guided" (Frankfurt 1988 175). Some decisions, therefore, bear not just on a specific situation but have consequences for future situations as they can partly determine the processes that contribute to those future decisions.[371]

The upshot of this analysis is that an agent can try to resist "doing what comes naturally" by putting constraints on those processes that would otherwise determine in an unreflected manner his actions and thus sculpt the space of actions that remain open for him to do in a given situation.[372] The functions of these constraints are twofold: they create or contribute to the coherence of an agent's intentional actions and they provide a 'reflexive or hierarchical structure' between his desires and consequently to his identity, which does also contribute to the coherence of his actions. Indeed, Frankfurt applies elsewhere the notion of 'person' to this constellation. A person, he writes, is characterized by taking upon himself constraints that are not just limiting his thought and language, but also the "choices he can make" (Frankfurt 1999 113).[373]In sum, the process of the person's engagement with his own 'psychic characteristics' is comparable to what we refer to as the agent's sculpting process with respect to his space of actions: in both cases constraints on the available options are established such that a more or less coherent pattern of performances emerges.[374]

---

[371] Indeed, an assumption of Frankfurt appears to be that our decisions somehow influence the neural and cognitive processes that are involved in future decision making. He does not explicitly address this answer in the contexts discussed here, but others do. Taking up Frankfurt's influential notion of guidance control – discussed above in section III.2.1 - Fischer and Ravizza argue that in fact also the mechanisms underlying our decision making must be moderately reasons-responsive (Fischer and Ravizza 1998). Contributing to this discussion, Frankfurt articulates a very loose sense of 'reason' when he even ascribes reasons – but not beliefs - to insects that try to escape a predator (Frankfurt 2002a). Fischer agrees with that loose sense, writing that: "[a]n organism – any organism – can have reasons insofar as he or she can have interests or a "stake" in something" (Fischer 2004 149).

[372] Frankfurt has a keen interest in the psychological plausability of what a philosophical analysis suggests that an agent should mentally accomplish. For example, he notes that there may be a trade-off between the size of an agent's enlarged space of actions and his sense of identity: "[t]he task of evaluating and ranking a considerably enlarged number of alternatives may be too much for him; it may overload his capacity to make decisions firmly grounded in a steady appreciation of what he really values and desires" (Frankfurt 1999 109). Having established a stable set of constraints does, from this perspective, not so much restrict an agent's identity in a negative sense but rather supports it.

[373] With the importance attached to the person and its structure being determined by the constraints he has taken upon himself, Frankfurt's account offers little room for agents who try to justify an action retrospectively by adjusting their second-order volitions such that their action now complies with those. Such an agent, we would argue, has not determined the structure of his will and can be compared to the 'wanton' who is not moved by his will either (Frankfurt 1971).

[374] In an earlier text, his focus was less on the notion of the person but rather more on guidance and control. In that context, too, the emphasis is on the fact that an agent can avoid the performance of undesirable proximal intentions by constraining these, that is to: "replace the liberty of anarchic impulsive behavior with the autonomy of being under his own control" (Frankfurt 1988 175).

However, what remains to be shown is: why should an agent develop these characteristics at all, what good are these constraints? Indeed, Bratman has argued that a purely hierarchical account might be subject to a threat of circularity, it being that an agent might support some of his intentions to act by referring to his identity as a person, even though this identity is found to partly consist of precisely such selected intentions. Adding a temporal or dynamical component to this hierarchical account – since an agent is not a 'time-slice agent' - it can be pointed out that an agent cannot but develop higher-order policies that allow him to self-govern his actions. Only by doing so can he at least try to avoid undesirable incoherencies and inconsistencies in his actions, for example by blocking an action in an exceptional situation, as will be discussed in the next section.[375] Nonetheless, these policies are not just based upon mere instrumental means-end reasoning but involve the agent's setting an end to his reasoning that is not instrumental in nature (Bratman 2002). Because of the constraints of his temporally extended agency, an agent must develop the constraints mentioned earlier – both contributing to the sculpting process as a whole. In what follows, we will further consider the constraints that pertain to the level of proximal intentions and – subsequently - their implementations.

3.1.2 Proximal intentions and blocking habitual action

As noted earlier, intentions occupy multiple positions in the agent's overall psychology with regard to his actions, ranging from prompters to terminators of deliberation, and from initiators to guidances of action. Corresponding with these positions are different functional properties, corresponding with the roles these intentions play in the complex and dynamic processes involved in his agency. An important role of Bratman's present-directed intentions – to which we refer as proximal intentions – is their being directed at the present or proximal situation, recognizing it as an appropriate situation for carrying out a specific intention: "[t]o have a present-directed intention to A, I must see that *now* is the time for action" (Bratman 1987 182, note 8, italics in original). Seeing that the present situation enables carrying out an intention is often, as we have learnt above, a matter of pattern recognition for which motor intentions are responsible. However, the recognition of a suitable affordance that is associated with a particular motor representation may trigger behavior that is not intentional but merely a habit – it may indeed even be contrary to an agent's current

[375] Obviously, with distal intentions in place, an agent is still not invulnerable for inconsistencies. First, it is improbable that an agent can specify all proximal – let alone motor – intentions to such an extent that he will never realize that in a given situation not all distal intentions can consistently be carried further. Second, as distal intentions often need time to be articulated and further specified, potential inconsistenties to be solved often emerge only at a later point.

intentions, at whatever level of specificity.[376] So how are the two different and what different cognitive processes are involved?

Bratman recognizes this risk of lapsing into habitual action in a situation where the outcome of that action would be at odds with our long-term or distal intention. Appreciating, like Frankfurt did, the limitations in cognitive resources and time for an agent to consider all available action options in a given situation, still Bratman draws attention to the important fact that an agent may at times be forced to 'block' the application of a general intention. He mentions the example of an emergency situation, in which case a car driver is forced not to buckle up, even though it does not mean that he reconsiders or abandons the general intention to always buckle up (Bratman 1987 88-89). The difference between a habitual response to a circumstantial trigger and an intentional response in Bratman's sense, is precisely this defeasibility of the latter as can be observed when it is being blocked in extraordinary cases. In such cases we can observe that there are several constraints or rules at work, interacting with each other in rather complex ways.

Proximal intentions are in a way a focal point of such interaction of multiple constraints. For Bratman has assigned to proximal intentions several functional roles. Proximal intentions play a role in initiating an action at a particular moment in a given situation, without depending upon a careful deliberation of the pro's and contra's of a particular action since that deliberation has been part of the distal intention against the background of which a proximal intention is formed.[377] However, anchoring a distal intention by way of a proximal action in a given situation does often require adjustments of the latter, without adjustments of the former. For example, adjusting the proximal intention to a changing situation involves 'temporal updating' it so that the 'now' remains appropriate (Bratman 1987 56).

Such updating of a proximal intention in order to anchor a distal intention appropriately or to block its application if necessary is influenced by some further constraints to which we will turn now.

---

[376] In footnote 329, we've referred to the fact that habits have a fixed nature and do not result from a competition between potential action responses in a given situation. See (Graybiel 2008) for further clarification of habits.

[377] Bratman refers to the 'hybrid character' that many intentions and action have with respect to their being deliberative, since many are only deliberative because of their being part of a comprehensive plan which is the result of deliberation and is itself not reconsidered in a given situation (Bratman 1987 30). Similarly, the standards or criteria to which such intentions and action should conform are only derived from such a plan or distal intention.

3.1.3 Proximal intentions and constraints for anchoring an action

Not just prompting but also exercising some control on an agent's temporally extended action, the question is whether particular standards or criteria are involved that may bring the agent to abort or instead release his action. Bratman argues that indeed such critera are in place and these are the criteria that we also apply when judging the rationality of an agent's explanation of his action in terms of his intentions. Such judgments are not only made regarding the distal intentions of an agent but also with regard to his proximal intentions. This argument results in the formulation of an 'intention-action principle': "the present-directed intention to A and the resulting action of intentionally A-ing are too tightly connected for us to praise the agent as rational for the former and yet not praise her as rational for the latter. This is because the intention and action are not separately controlled by the agent, but rather the agent's control of her action goes by way of her intention" (Bratman 1987 55). What standards or criteria are applied for this control process?

Bratman presents us with two different norms or constraints that – explicitly or implicitly – should apply to an agent's planning, if he is to optimally fulfill his intentions and plans.[378] These constraints respond to the fact that an action is never instantaneous nor isolated but interacting with other actions, intentions and beliefs – in ways, however, that are not completely transparent to the agent. The coordinating role of the agent's intentions and plans consist partly in putting constraints in place on the processes that eventuate in action and this holds not just for distal intentions but even for proximal intentions. The first constraint demands that a plan (or plans) should be not self-refuting but consistent, while the second constraint demands for means-end coherence of the plan. A planning agent is one who does take these two constraints into account, one way or another: "The recognition of these demands helps distinguish intentions and plans, on the one hand, from ordinary desires and valuations, on the other" (Bratman 1987 32). Given this analysis, an articulation of the constraints is relevant.[379]

Take the twofold consistency of an action plan, needed for effectively carrying out an intention. The demand for an intention's internal consistency is obvious as it is inconsistent if a singer plans to be silent and to sing a line simultaneously. Such

[378] Quite consistently, Bratman formulated these demands, constraints or standards in his early 1981 article on means-ends reasoning (Bratman 1981), in his 1987 book (Bratman 1987), in (Bratman 1992a), his 2006 book (Bratman 2006b) and in his (Bratman 2009b).

[379] So it is foremost due to functional rather than moral considerations that we can us expect this complexity of the agent's cognitive processes with regard to proximal intentions. Indeed, Bratman argues that in the end the functionality of these proceses derives largely from "our interest in getting what we want" (Bratman 1981 262).

consistency forces an agent to also take his beliefs into account (while assuming that these are true) as he is constructing a representation of reality: "my plans should fit together with my beliefs into a consistent conception of the future" (Bratman 1981 259). Such a construction is difficult as it is complicated to account for the future since this is partly shaped by the agent's own intentions and plans – even in the case of a proximal intention and the immediate future that may result from it.[380]

The second constraint an agent should fulfill is a result of the fact that an action is usually not only temporally extended but also involves a hierarchy of steps or means that contribute to performing it.[381] Correspondingly, an agent's action plan needs to consist also of "subplans concerning means, preliminary steps, and relatively specific course of action, subplans at least as extensive as I believe are now required to do what I plan" (Bratman 1987 31). The specification of action, which the framework of the intentional cascade ascribes to proximal intentions, is at stake here. To begin with, it requires the agent to develop such a specification at all. He cannot pause with the mere formation of an action goal but should indeed proceed to specify means to realize the action – for example specifying whether to use precision grip or power grip while picking up a cross or sword. Moreover, these means should be included in his intention or plan if it is to count as such: we would consider it irrational for an agent to intend to reach a certain end yet not to intend executing a means which he believes to be necessary for reaching it (Bratman 2009b). Irrationality would in this case amount to dysfunctionality since an intention can not be realized without the intention of realizing its means. Determining the level of detail of this specification of the means, however, is difficult.

Clearly, it is not necessary for an agent to – perhaps implicitly – take every possible future situation in consideration, as some of these situations are implausible or unlikely to happen. Similarly, he is not required to specify in advance all minor

---

[380] Kolodny discusses the 'predictive significance' of intention (a term introduced by Scanlon) with respect to the future in connection with the constraints on practical reasoning. Kolodny underlines that consistency demands can make a rational agent to develop further intentions if these are related to actions that will facilitate the satisfaction of an earlier intention, which is different from mere means-end reasoning (Kolodny 2008).

[381] Bratman underlines that a planning agent is not a time-slice being, partially because his agency is extended in time (Bratman 2006b). If the agent is to self-govern, then he needs to realize that self-governance is not a time-slice phenomenon, too. Instead, it requires temporally extended planning with an important role for means-end coherence and limited room for reconsideration of his plans (Bratman 2007).

[382] Obviously, details of the muscle movements that are involved are even impossible to specify. Although most authors agree about this, there is still some debate about the question whether experts can or cannot articulate and verbally express details of their expertise. We've touched upon that debate in the previous Part. There we did i.a. refer to the implict and explicit stages of learning and development presented in (Karmiloff-Smith 1992), which presents a 'representational redescription' account of empirical research of expertise, including a final stage in which expertise can be made explicit. This account was contrasted with an analytical one that denied such explicitability of expertise (Dreyfus 2004).

details that will be involved in performing the action.[382] Depending on the agent's expertise, habits and skills, he must specify his intention more into detail or he can leave it up to the moment of performance and then rely on these for appropriately performing the action (Bratman 1987). For example, for a beginner without such expertise and habits, it is necessary to specify more into detail how he is going to perform next his canzonetta for Donna Elvira's maid, while an expert may trust his already stored proximal intentions to include the necessary specifications required at the time. This expert, though, may have to block and re-adjust his proximal intentions in a new and divergent stage direction. Clearly, anchoring his performance in that situation consistently and coherently may require him to specify his behavior, singing and instrument use completely different – not just during his preparations but also on-stage during his performance. A continuous complex interaction of established and newly formed intentions is the result, pointing towards the peculiar position of proximate intentions.

3.1.4 Proximal intentions and their peculiar position in the agent's psychology

The role played by an agent's expertise, habits and skills – which were also associated with motor intentions earlier - confirms Pacherie's observation that proximal intentions occupy an intermediate position in the framework (Pacherie 2008). Although proximal intentions can provide some amount of conscious guidance to the ongoing action, they are assisted in carrying out this complex task by the presence of motor intentions that can to some extent relieve it (Pacherie 2006). As a result, proximal intentions inherit constraints of a sculpted space of actions from both motor and distal intentions.[383]That they still have a specific role of their own is most obvious in a situation in which the agent demonstrates the 'defeasibility of general policies' by not applying a policy to the particular case, for example when otherwise a breach of some of his other constraints of consistency and coherence might occur (Bratman 1987). Our expert singer, for example, must be able during a extraordinary Don Giovanni stage direction to inhibit commonly juvenile behavior as his distal intention of complying with a director requires him now to act otherwise.

Let us try to shed some light on this peculiar position of proximal intentions. In some sense they appear to function like kludges that have been established in a responsible mechanism. Yet in another sense and unlike kludges, proximal intentions

[383] Hobson explicitly doubts whether Bratman's present-directed (proximal) intentions as such exist. Although he is correctly pointing out the priority of future-directed (distal) intentions, he overlooks the importance of the function of proximal intentions in anchoring the latter in specific situations and – at times - blocking their application (Holton 1999 246).

are also governed by constraints that are associated with conscious deliberation. What kind of explanation can be given for this peculiar nature of proximal intentions? To be sure, intentions and plans do not fulfill their roles in isolation from other components in the psychology of an agent.[384] Irrespective of the fact that we can subject them to those rational constraints as were presented above, they are also related to constraints that depend upon the physical and psychological beings that agents are. Indeed, the rational constraints themselves are partly due to our physical and psychological properties and limitations, which disallow us to perform two contradictory actions simultaneously, for example. Philosophical analysis aims partly to clarify precisely the interdependence that exists between our psychological structures – studied in psychology and neuroscience - and the structures of our thought and action (Bratman 2009a).[385]

Indeed, the present analysis of the process of sculpting a space of action aims to contribute to such a clarification, as well. It more specifically focuses on the properties of this sculpting process, which are partly dependent upon the constraints that have their source in practical deliberations that do at times determine the actions an agent performs. For another part, the constraints stem from the embodied and cognitive structures that constitute the agent and which also constrain the algorithmic and neural implementations of these deliberations.[386] Integral to these various constraints is the requirement that the space of actions does not remain completely fluid but gains a profile that is rather stable. By adding generatively entrenched properties to such a

---

[384] Intentions can be recognized as distinct psychological elements in the philosophy of mind, making their reason-giving status in normative philosophy only derivative, Bratman argues (Bratman 1981 263). There is some analogy in this view to Frankfurt's admission that it may be possible that an immoral life can be valuable and desirable for the person that lives it, who may not have convincing reasons to change it as long as it is coherent and consistent. Frankfurt explains: "the value to Hitler of living the life he chose would have been damaged by the immorality of that life only if morality was something that Hitler actually cared about, or if the immorality of his life somehow had a damaging effect on other matters that he cared about" (Frankfurt 2002b 248).

[385] Early on, Bratman committed to a functionalist approach according to which inputs are connected through psychological processes and activities to outputs like actions. There are regularities involved in these processes and activities, for which intentions, beliefs and desires can be responsible. An intention is then to be understood as a 'distinctive attitude' (Bratman 1987). In line with this account, Bratman much later acknowledged in his contribution to a symposium on consciousness that a specific intention needs not to be engaged in conscious thinking continuously. He referred to the Freudian idea of unconscious intentions that still have specific content, like 'sleep with your mother'. Consciousness would than be a higher-order – relational - phenomenon, that may or may not be at stake with regard to a specific intention at a specific moment (Bratman 2006a).

[386] As noted in Part I, such a multicausal account of human action is not new. Aristotle has offered a first account along such lines, famously contending that: "Thus every action must be due to one or other of seven causes: chance, nature, compulsion, habit, reasoning, anger, or appetite" (Rhetorics 1369 a 5-6). Our scrutiny of the role of a sculpted space of actions is partly inspired by the emphasis that Aristotle put upon the role of habit in action – including moral action in his Ethics.

sculpted space the agent will be faster and more flexible in his responsive actions, with decreased recruitment of resources in their performance. For the execution of his own complex actions as much as for joint or collective actions of a certain complexity, this sculpted space is quite advantageous (Keestra 2012).

Moreover, an important lesson that must be drawn from the consideration of this interdependency and the constraints that are associated with it is that an agent should basically not modify continuously his intentions and plans but should remain largely committed to them. Would an agent instead constantly reconsider and modify his plans, he would be seriously impeded in completing an intentional action at all (Bratman 2006b). This is partly so because generally a plan has ramifications and is temporally extended, rendering it likely that a modification will turn out to be counterproductive as it runs against earlier phases of its execution.[387] Apart from this argument about the irrationality of continuous reconsideration of one's intentions, Bratman mentions another reason why such reconsideration is disadvantageous to an agent – yet this reason is of a naturalistic nature and testifies to his ambition of developing an account of agency that fulfills several roles.[388]

If an agent aims to avoid counterproductivity as a result of his reconsidering his intentions, he is forced to carefully scrutinize the modification and its consequences which "is an activity that uses up time and other limited resources; while engaged in reconsideration I am unable to do other valuable things" (Bratman 1992a 6).[389]Apparently, this costliness has to be taken into consideration as well as it also affects the agent's adequate actions. When modifying his plans and intentions, he must also devote cognitive resources to updating his proximal intentions, impeding his ongoing performance.[390] In contrast to investing these resources in case of

---

[387] Given the temporal extendedness of any action, it is unsurprising that a methodological priority is given to distal intentions, as they most distinctly take future consequences and situations of action into account (Bratman 1984).

[388] In an early article in which decision-theoretic and AI approaches to rational agents are being combined, Bratman and others take resource boundedness explicitly into account. In that context, the subjective expected utility is taken to be 'a function of the agent's beliefs and desires.' Later, however, the cost of the process of deliberation itself is also mentioned, which seems relatively independent from the agent's beliefs and desires (Bratman, Israel et al. 1988).

[389] Bratman is aware of the risk involved in the recommendation not to constantly reconsider one's intentional structure, for he suggests that it may turn out to be necessary to formulate a historical condition which explicitly excludes the influence of extreme cases of manipulation, brainwashing and the like from it (Bratman 2003).

[390] In his project of 'creature construction' to which we referred earlier, at a relative early stage Bratman observes that there are 'substantial pressures for mechanisms' that "support coordination, intrapersonal and interpersonal, in ways compatible with these limits" (Bratman 2006b 53). Only later do capacities for deliberation about these come into play, which yield additional benefits but still need not always to be employed.

reconsideration, intentional action can normally rely upon dispositions that are the result of our having established a ramified structure of intentions.[391] Such dispositions, then, comply with the constraints mentioned earlier and are the result of the agent's long-term fostering these constraints, even though they are no longer dependent upon the reasoning, deliberating and conscious consideration that were at some point associated with them.[392] An agent's established structure of intentions and the dispositions that can emerge from it together does contribute to his constrained or sculpted space of actions, in our words, and thus facilitates the agent's cost-effective, fast, flexible and intentional action in most cases.[393]

3.2 Proximal intentions and cognitive mechanisms that determine anchored actions

If an expert singer has developed a distal intention to perform Saint François not in a solemn but in a nearly heroic manner though not too juvenile, he needs to recognize when is the right situation to sing a particular phrase in a way that fits that intention. Perhaps not all parts of the huge score are particularly apt for his peculiar interpretation, for instance as his vocal interactions with birds do not offer an appropriate situation to display heroism. Such an agent may have several distal intentions, all lingering simultaneously in the background and waiting for the appropriate occasion to be realized, while other must be blocked given the conflicting current stage direction. Proximal intentions are needed to navigate in a particular situation, helping him to anchor one or more distal intentions in that situation. These proximal intentions are subject to constraints of consistency and means-ends coherence, which make

[391] Indeed, following up on Bratman's theory of planning agency, Pollard argues for a more prominent role of habits – now not understood as mere reflex-like habitual actions. According to his account, habits are not so much dependent upon beliefs, desires, intentions and reasons but contribute to actions in a rather more embodied and embedded way (Pollard 2006). Elsewhere, such identification of an agent's ability to act with his having appropriate dispositions – as a result of having acquired habits and skills – is labelled 'New Dispositionalism' (Di Nucci 2011).

[392] Kolodny argues that it is not uncommon to praise an agent who has certain unconscious dispositions in the absence of conscious deliberation of reasons for an action. These dispositions may make 'believers and intenders' "sensitive to these reasons, either via their beliefs about them, or via unconscious mechanisms" (Kolodny 2008 390). When reasons offer constraints on actions, it is plausible that they allow processing by unconscious mechanisms.

[393] In his contribution to a symposium on consciousness, Bratman argued that a specific intention needs not to be engaged in conscious thinking continuously. He referred to the Freudian idea of unconscious intentions that still have specific content, like 'sleep with your mother'. Consciousness would than be a higher-order or relational phenomenon, that may or may not be at stake with regard to a specific intention at a specific moment (Bratman 2006a). Apparently Bratman agrees that intentional contents may contribute to an agent's actions implicitly. Such an implicit intention may still become explicitly aware to the agent. Similarly, it may have originally been consciously made. This argument also reflects his interest in developing an account of agency that is not just theoretically sound but also empirically plausible.

them different from motor intentions, which could be simply triggered by perceived affordances. Given an agent's commitment to more comprehensive action plans, proximal intentions therefore play a crucial role with their function of anchoring and specifying action. The intentional cascade framework allows us to discuss how representations of an intention to act at several levels of specificity and involving different elements are involved in the complex cognitive processes that determine an action. In the next sections we will consider whether cognitive (neuro-)scientific evidence concurs with the description of proximal intentions that we offered above. Before introducing some preliminary reflections about this, the next paragraph will remind the reader of some important lessons from our earlier discussion of empirical evidence regarding motor intentions.

Two lessons had great relevance and concerned the phases in skill learning and the modifying structure of the representations involved in this. We scrutinized empirical evidence with regard to skill learning by focusing on research of motor representations, which are capable of implicitly integrating both environmental information and options for motor action and enabling intentional action on a relatively short-term scale (Jeannerod 1997). This research provided as a first lesson that skill learning or growth of expertise corresponds with two phases that affect motor intentions: an early phase characterized by increasing efficiency of neural activations when skilled behavior is performed and a second phase, in which this behavior is strongly associated with additional activations in other neural areas, facilitating access to other representations or programs related to the skill (Petersen, van Mier et al. 1998). Second, the formation of these kludges during the process of developing a skill or form of expertise is associated with the development of chunks and templates that allow faster and more flexible processing of increasingly complex representations (Guida, Gobet et al. 2012). When we now turn to considering empirical evidence bearing relevance for proximal intentions, we may ask whether we will observe similar varieties in processing types and in forms of representations. Given the fact that proximal intentions are not restricted to processes that are by definition very fast and that escape consciousness, these varieties may present themselves in a different way. Indeed, instead of the distinction between two phases of skill learning, we will below be dealing again with a dual-process theory – similar to the dual-process theories that we've discussed in chapter II.3 ff. This turn to a dual-process theory is not made by Pacherie, who made a different choice in the present context. So why did we choose otherwise?

Pacherie does refer to some evidence with regard to proximal intentions. Accounting for the complexity of the plans involved in these, she mentions the intentional schema theory, which is built largely upon developmental, ethological and psychological evidence. Acknowledging that this theory aims to explain intentional social interactions, her main interest is in the way that intentional schemas integrate perceptual and action related information while being hierarchically organized (Barresi and Moore 1996). Pacherie furthermore has added to her intentional cascade also forward and inverse models of motor actions determined by the intentions that figure at three different levels of the cascade. These models are used for the prediction, control and correction of actions, by being involved in cognitive processes which amount to their comparison. Like the intentional schemas we just referred to, these motor action models integrate information representing the situation with information that represents the action itself (Pacherie 2008).[394] Given that Pacherie's own account remains relatively abstract with regard to the information involved in proximal intentions and their implementation – partly because her interest is mainly in the phenomenology and experience of action and less in its determination – and given the fact that the intentional schema theory has also a relatively small empirical basis, we will below discuss whether cognitive neuroscientific insights allow us to be more specific about proximal intentions. Doing so, we will focus on a particular cognitive neuroscientific account, that offers the ingredients that are required once we intend to explain the complex functional properties of proximal intentions.

With proximal intentions, the explanatory task will be more complex than for motor intentions. Given the dual function of proximal intentions to both anchor and specify an intention in a given situation but also to block its performance in exceptional cases (Bratman 1987), their explanation must accordingly cover such divergent properties.[395] Moreover, we also noted that a person's identity and the hierarchical structure of his intentions and the constraints these put upon his intentional actions is at stake in this context (Bratman 2006b). Facilitating our explanation, however, will be the fact that for our explanation of these functional properties we will also employ the preceding explanation of motor intentions, as these are indeed generatively entreched and being employed in proximal intentions, as well. Integrating elements of the explanation of motor intentions, we might expect the explanation of proximal intentions to be more complex both in terms of processing and of representations. Let us first take a quick

---

[394] Elsewhere Pacherie adds to her model a distinction between two different forms of control – following an analysis offered in (Buekens, Vanmechelen et al. 2001) – namely: tracking control and collateral control. Tracking control is engaged for flexibly adjusting the motor action in order to succesfully reach the goal, collateral control is controlling for undesirable side effects (Pacherie 2006). Note that the latter form of control could amount to blocking the execution of the action, which is one of the functions of proximal intentions according to Bratman (Bratman 1987).

[395] Concurring with Bratman's intention-action principle which holds that an agent controls his actions via his intentions (Bratman 1987), many empirical studies of the control and determination of action are comparably interested in the representations and reasoning that are involved in the agent's intentions.

look at some notions that have been proposed to account for some properties we associate with proximal intentions.

During the last century and particularly its last half, researchers have developed constructs that enabled them to explain the quite divergent functional properties of intentional actions, the changes that occur to these due to development and learning, and the errors or deviations that do obtain in exceptional cases. Generally speaking, these constructs were primarily presented in terms of the representations and cognitive strategies involved, with their neural implementations being explained only more recently. An early example is the 'schema' construct. It was put central by Bartlett in 1932 – who actually preferred the term 'active developing patterns' to the word 'schema' - explaining how an individual always integrates novel information into representations that have developed from past experiences, instead of his merely assembling separate memories (Bartlett 1995 [1932]).[396] Such hierarchical structure emerged not only from empirical but also from computational studies. Indeed, based as they were upon not only psychological but also on computational and simulation studies, the term 'scripts' was suggested for those higher level representations of temporally extended actions (Schank 1980).

These knowledge structures have been not only applied to cognitive processes but also to other domains, like Piaget and others did when they used schemas to account for developments in behavior, language and thought (Arbib 2003b). Indeed, the study of skill learning did equally reveal the importance of hierarchically structured representations, with 'plans' becoming increasingly stable and governing over time complex habitual actions in an automatized fashion (Miller, Galanter et al. 1960) Moreover, this employment of hierarchically structured schemas – implicitly in most cases – could not only explain the structure of both behavior and language, but also explain exceptional cases or disorders in which the serial order of actions is affected, for example (Lashley 1951).

However, research did not only demonstrate how the representations involved in various forms of intentional action are complex and hierarchically structured, it did also suggest that more than just a single cognitive process might be involved. Indeed, such representations are considered as frameworks which are the result of the integration of several sources of information and which are also employed by different

[396] Interestingly, Bartlett explicitly contended that schemata also include social and cultural information. This social aspect, though, has been left out of the schema concept when it was taken up in cognitive science some decades later only to reenter the research more recently (McVee, Dunsmore et al. 2005). Indeed, in our analysis of distal intentions below, which refer to comprehensive social and cultural influences will be present as well.

processes simultaneously (Minsky 1975). An example of such parallel employment is when schemas are not only being used in the initiation and determination of actions but also play a role in the ongoing control and correction of these (Arbib 1981). Or, as another computational model accounting for psychological evidence suggests, information is integrated that has been retrieved from both long-term and working memory, which underlies both the flexibility and increasing stability of cognitive, language and behavioral skills (Anderson 1983).

What can be gathered from this short glimpse of some relevant theories is that explanations for the complex functional properties of expert behavior are indeed striving to combine insights regarding both representations with their complex structures and different processes with potential interactions. We have chosen to organize our discussion of a plausible explanation of our account of an agent's sculpted space of action by focusing on an influential theory and corresponding model of action determination that has been built upon the notions that we just mentioned, has been tested against empirical findings, and has accordingly been subject to proposed modifications. Norman and Shallice have developed a kind of dual-process theory of 'willed and automatic control of behavior', which has been primarily developed as a computational model for every day actions and for action errors or disorganization that can be seen in healthy subjects and patients. Based upon clinical and other observations, the theory meant to "account for the ability of some action sequences to run themselves off automatically, without conscious control or attentional resources, yet to be modulated by deliberate conscious control when necessary" (Norman and Shallice 1986 378).[397]

Taking up several features of the constructs that were mentioned just before, actions are in this theory, too, taken to be hierarchically structured, with lower level motor schemas[398] containing sensory-motor mappings that determine muscular movements and higher level scripts determining the ordered and adequate performance of such motor schemas when performing a complex and temporally extended action (Cooper

---

[397] In our discussion we will not just rely on the original formulation in (Norman and Shallice 1986) but refer to the updated and expanded version of the theory as for example presented in (Cooper and Shallice 2000 ; Cooper and Shallice 2006).

[398] The notion of motor schema is similar to that of Jeannerod's motor representation, referred to in our sections on motor intentions (Jeannerod 1994): it integrates both information about affordances and about specific motor actions. In the current context, however, the additional question is how several such motor schemas are carried out in an appropriate order such that the performance of a complex action is enabled by them (Cooper and Shallice 2000).

[399] Indeed, the updated and implemented version of Norman & Shallice's model does incorporate the hierarchical analysis of everyday action and disorganization errors as presented in (Schwartz, Montgomery et al. 1995 ; Schwartz, Reed et al. 1991). As a result, the updated model distinguishes five levels of action control, ranging from, for example, the upper level of 'morning routine' to the lowest level of 'take cream container', belonging to the sub-routine of making coffee as a part of breakfast (Cooper and Shallice 2000).

and Shallice 2000).[399]Such representations are the result of learning and experience and include information about relevant environmental trigger conditions like objects that can be manipulated, about goals or end states to be reached, and about the relevant motor behaviors that might realize such end states. The result of the complexity of these representations is that an action representation can be activated via a variety of triggers, as our expert singer's representation for singing might be activated via the perception of a mandoline or of Zerlina, for example.

Specific for this theory is the assumption that such action representations are established with learning and experience, and then figure in an 'interactive activation network'. The latter implies that action representations can be considered as nodes in a complex network, with each node having a variable activation level that is determined by many different factors, including environmental triggers, motor effectors, and other ongoing processes. Moreover, each action representation can also influence the activation levels of other representations, for example by increasing the activation level of its own lower-level, component actions representations and – conversely - by inhibiting the activation of representations of competing actions. A routine action is carried out once a particular representation is activated beyond a given treshold and has outcompeted rival ones that have become activated as well.[388] As a result, well-learned actions can accordingly be performed without any influence of conscious or deliberate control, simply as a result of a change in a representation's activation level (Shallice 1988).

Apart from the 'contention scheduling' (CS) that is responsible for automatic control of actions, this model describes a 'supervisory attentional system' (SAS) that is responsible for deliberate action control. According to the model, routine action and cognition would not require such control, leaving top-down control necessary only "if error correction and planning have to be performed, if the situation is novel, or temptation must be overcome" (Norman and Shallice 1986 382). in novel tasks or in complex tasks dorsolateral prefrontal cortex activation is increased. This is interpreted as activation associated with the acquisition or generation of rules that help to select a response (Crescentini, Seyed-Allaei et al. 2011 ; Frith, Friston et al. 1991).[389] In its original formulation, this SAS exerts its effects by a top-down process that amounts to focusing attention on a particular action representation or relevant features of it.

388 This process of the selection of an action representation under specific conditions is described more in detail as 'contention scheduling' by the authors and consists mainly of activation and inhibition of potentially performed representations (Norman and Shallice 1986). We will only present this process to the extent that is necessary here.

389 Indeed, motivated partly by this CS/SAS model, a study investigated subjects' employment of rules that constrain the search space in a response task. In that context, the authors use the phrase 'sculpting the response space', which has inspired the title of this dissertation (Fletcher, Shallice et al. 2000 ; Frith 2000).

In doing so, it primes or biases the corresponding action representation and thus can modulate the CS process. Further elaborations have brought about a subdivision of SAS processes, all of which are still considered to modulate the activation levels of action representations and in that sense cohere as a supervisory system.[390] Different supervisory subprocesses that are dinstuighished are the spontaneous generation of an action representation, the specification of an action in order to solve a particular probem or the retrieval from memory of a particular action representation (Shallice and Burgess 1996). As these processes cannot directly influence the performance of an action but only indirectly by increasing or decreasing activation levels of action representations, top-down influences on an agent's actions are always competing with the other processes that determine his eventual performances.

With these two processes of 'contention scheduling' and 'supervisory attentional systems' in place, let us consider whether they can provide some more insights in the properties of proximal intentions that philosophical analyses have presented to us. Moreover, they may perhaps add to these analyses, while we may, conversely, make suggestions for further developments of the cognitive neuroscientific theory and model.

3.2.1 Processes for the resolution of conflicts between action options

In section III.3. we reminded the reader of the possibility that a single opera scene might offer many affordances for action, which suggests that more than just a single motor intention might be activated. How would our expert singer be capable of fast and adequately selecting only one motor intention for performance? Even more problematic seemed the situation in which he would have to block the performance of a habitual motor intention in an exceptional situation, as when a stage director requires a senior Don Giovanni or a heroic Saint François. Finally, we asked how coherence over time between intentions was created, given that motor intentions did not appear to provide for this.

We learnt from our subsequent discussion of arguments by Frankfurt and Bratman that selecting an option for action from the many options available in any given situation is a complex task which asks for processes and choices that far transgress the

[390] In singing at least four processes have been distinguished on the basis of collected evidence: auditory perception, a decision making process for retrieving and selecting an appropriate representation, the execution of the corresponding motor plan and finally the evaluation of the outcome of it – the sung tone or melody (Hutchins and Peretz 2011). Experiments with expert singers show their representations to be hierarchically structured, depending on their expertise with musical structure and performance (Zurbriggen, Fontenot et al. 2006). The complexity of these representations can be more complex than those that have been discussed in the context of motor intentions, which are limited in temporal extension and play their role mostly implicitly.

boundaries of any particular situation. Solving this task and resolving the potential conflict between action options requires an agent to organize and coordinate his intentions for action such that a coherent structure determining these intentions emerges. This coherent structure consists of 'constraints' that an agent takes upon himself and that limit the space of options that are open for him to perform (Frankfurt 1999). Further elaborating the coherence of this structure and analyzing how it affects an agent's actions in various situations, Bratman has noted a tension between two requirements to which an agent has to answer and which befall to proximal intentions to fulfill. On the one hand, an agent should refrain from costly and counterproductive reconsideration of the coherent structure of his comprehensive action plans, while also being prepared to block the performance of an action that belongs to such a plan, if there is a higher-level or distal intention that demands such a deviant response in a given – emergency, for example - situation (Bratman 1992a).

The dual-process CS/SAS model initiated by (Norman and Shallice 1986) provides several ways in which an agent can comply with the requirements that the analyses have presented – implicitly and explicitly. Did a motor intention integrate specific environmental information and a motor response option, the action representations involved in the current model can contain more information. With that, they are potential implementations of the proximal intentions discussed above. As a result of learning and experience, the scripts and schemas included in this model figure in comprehensive hierarchies of actions and also determine sequences of actions. Given the 'interactive activation' that determines the connections within these hierarchies and which is to a large extent the result of experience, there are continuously many vertical as well as horizontal interactions that contribute to an agent's sculpted space of actions. Vertical interactions obtain when a comprehensive action script has been initiated, for example by singing a Don Giovanni performance. In that case, the activation levels of a set of component action representations at multiple hierarchical levels is being increased. Horizontal interactions also occur, as when the activation level of competing representations is being inhibited or decreased at the same time (Cooper and Shallice 2000).

With such hierarchical activation patterns established over time, our expert singer will usually be able to perform coherently once the appropriate comprehensive action representation has been activated. One consequence of this is that motor intentions that do not belong to this representation will have less chances of becoming activated: in a given situation both irrelevant environmental triggers and inappropriate motor representations are normally less activated or sometimes even inhibited (Cooper and Shallice 2000).[391] When Don Giovanni is trying to seduce Zerlina, for example, it is most unlikely that he will use the scolding voice with which he will later answer Donna Elvira. Even more unlikely is that he will meditate upon a cross when having his dinner near the Commendatore's grave as our singer might do when playing Saint François or that he will accidentally pick a handbow from the wall as if he was singing Guillaume Tell. Indeed, with the addition of extra pre- and postconditions to the hierarchically organized component actions, the 'contention scheduling' model has received extra features that support his appropriate performance. According to these conditions, an action (component) will be activated by particular environmental features or by the completion of a preceding component. Conversely, the action will be inhibited if a particular goal or environmental situation obtains, even when the agent has not himself contributed to this to happen (Cooper 2002).

The study of deviant behaviors in healthy agents and in patients has influenced the development of this CS/SAS model and has also contributed to insights regarding the coherence and the flexibility of normal actions (Schwartz 2006 ; Schwartz, Montgomery et al. 1995). Results show that there are many ways in which actions can become disorganized, ommissions of action components occur, agents suffer from perseverations or unnecessary, repetitions of actions, misuse of objects for specific actions, and so on. These errors affect the 'intermediate level of organisation', as they happen in a time frame that would allow conscious control (Cooper and Shallice 2000), which distinguishes our proximal intentions from motor intentions. The current model offers several explanations for incoherencies visible in these errors. Since the activation level of a particular action representation is based upon the summation of activation levels that depends from excitation or inhibition through neighbouring action representations, from bottom-up excitation by environmental triggers and from top-down excitation via an activated higher level action representation, there are multiple ways in which this coherence and flexibility can be negatively affected. This is the case when the 'parameters' of interactions that are prevalent in a patient's CS mechanisms are set at a level that is so high or low that inappropriate excitation or inhibition obtains (Rumiati, Zanini et al. 2001). Obviously, the other side of the coin is that an agent's coherent acting can equally be explained via this model, even though it may be in terms different from the ranking or prioritizing of intentions for action that we discussed in section III.3.1.1.

[391] The theory about 'Structured Event Complexes' is developed to some extent in discussion with the CS/SAS model, although it has distinct ideas about neural implementations of the processes it refers to. Similar to both approaches, though, is the idea that interactive activations between component representations determine the selection of a hierarchically structured event during cognition or action (Grafman 1995). The SEC theory is also developed to account for the observation in healthy and normal agents that some SECs will be more rigid in their structure, whereas others are more flexible and allow ample adaptation to environmental conditions (Grafman 2006).

Most important, the compositions of the action representations' hierarchies and all activation levels involved are largely dependent upon the individual agent's development, learning and experience.[392] Given the Hebbian learning that is involved in these processes, all compositions and activation levels are obtained over longer periods of time.[393] The dynamics involved in these allow for many different influences, ranging from exposure to particular environmental triggers to deliberate choices that favor some over other actions and also provide quite some flexibility to each inidividual (Cooper and Shallice 2006).[394] For example, action observations and modelling studies suggest that the association between environmental triggers – comparable to affordances - and particular actions is a result of individual exploration and experience and allows adaptation to novel environments (Cooper and Glasspool 2001).[395] However, the question is whether the model allows the kinds of interaction between 'automatic and willed control of behavior' (Norman and Shallice 1986) necessary for prioritizing to occur: how can an agent deliberately control the automatic behavior that results from the CS mechanisms?

Obviously, where an individual agent will gather experiences and thus develop his interactive activation networks also in the absence of deliberate control, he has also the capability of supervisory control that allows him to exert some influence on these networks. In the initial formulation of the CS/SAS, it is the agent's 'attention to action'

---

[392] This is not to deny that there are many representations or components thereof that will be shared among individuals. Remember that in the previous Part we referred to a neuro-constructive account of development. Such an account emphasizes the brains 'embodiment' and 'ensocialment', which contributes to shared representations between individuals irrespective of their individual learning trajectories (Mareschal, Johnson et al. 2007). Before that, Karmiloff-Smith already emphasized the importance of innate prespecifications that facilitate the newborn's learning, challenging as she does both "Fodor's anti-constructivist nativism and Piaget's anti-nativist constructivism" (Karmiloff-Smith 1994 693). Apart from a shared social environment and stimuli that stem from this, these innate prespecifications also contribute to partly shared representations among agents from an early stage on.

[393] For example, switching to another task or course of action requires an agent to use his intentions for reconfiguring the activated action representations, which does require extra cognitive capabilities and resources. This explains why often agents persist in completing an initiated action (Goschke 2000). Task switching can be distinguished along a gradient which varies from exogenous control that is stimulus driven to endogenous control and has been associated with a posterior-anterior gradient along the prefrontal cortex (Kim, Johnson et al. 2011). As the CS/SAS model allows that the initiation of a comprehensive action activates simultaneously alternative component actions with their own trigger and goal conditions, expertise can accordingly indeed allow such task reconfigurations. However, activation through trigger conditions is powerful and it requires strong supervisory processing to override these (Norman and Shallice 1986).

[394] Obviously, there are many constraints at work that limit this flexibility. For example, house flies have to cope with physical constraints in their behavior that favor some behavioral sequences over others (Dawkins and Dawkins 1976), which holds to humans as well. Among others, however, cultural and individual constraints also determine the representations that drive human behavior (Grafman 1995).

[395] The template structure that we noticed to be present in motor intentions returns in the present context, too. The hierarchical action structures at stake here does not demand that representations are determined up to the lowest level, leaving room for flexibility in the motor movements to perform a certain action.

that can affect already existing preferences or help to respond to novel action situations (Norman and Shallice 1986). Later elaborations have distinguished between different supervisory processes that support control of actions by modulating activation levels such as spontaneous generation of an action representation, the specification of a goal-directed action or retrieving a particular action representation (Shallice and Burgess 1996). What he in fact is doing through such processes is priming targeted features of his own CS mechanisms.

Several strategies for controlling otherwise automatic actions have been discussed in section II.3.1.3 in the context of dual-process theories, which share commonalities with the current CS/SAS theory. A way of self-regulation would accordingly be for an agent to articulate a rather specific intention that he aims to implement in a situation in the near future, like responding to a particular environmental conditions with a specific action (Gollwitzer and Sheeran 2006). Using a converse strategy, he could prepare himself to inhibit a certain action by engaging in counterfactual thought which primes him for action options that are alternative to his routine ones by modulating his attention to otherwise unnoticed environmental triggers or his evaluation of undesirable action consequences (Galinsky and Moskowitz 2000). Indeed, part of the supervisory processes of the CS/SAS model are also processes that monitor and evaluate actions (Shallice and Burgess 1996).

Even though these supervisory or attentional processes are cognitively demanding and recruit resources that automatic action does not require, agents are capable of influencing their actions by employing them. Yet for effectuating lasting changes in his action routines, it would require him to realize long-term intentions to modulate these. On the other hand, by the time an agent has developed over time the action representation hierarchies and activation levels that conform to his intentions, 'contention scheduling' does allow him to act according to his intentions even in the absence of conscious control. This raises the question how blocking a routine action can take place, it being one of the roles that proximate intentions play.

3.2.2 Blocking a habitual action according to the CS/SAS model

As proximate intentions should among other functions allow an agent to block the application of a distal intention in a specific situation we should hope that hierarchical action representations do not rigidly require their comprehensive execution upon an agent from the moment of its initiation. In terms of Bratman's own example of not buckling up in a case of emergency (Bratman 1987), we would hope that the comprehensive script for driving, including as a component action one's buckling up, allows to be modified. Driving, for example, would then be subordinated to the superordinate action of bringing someone as quickly as possible to the hospital, which corresponds with adjustments like the following: component actions that are irrelevant for the superordinate action can be left out completely; objects that distract from the primary task should receive less attention; the physical well-being of the patient should receive some extra attention. As a result, the activation levels pertinent to component action representations, to environmental triggers, to action goals – or criteria for their fulfillment – require modulation. Such modulations can be hard to reach, since it may be difficult for an experienced driver to omit buckling up before driving. A comparable flexibility for our expert singer would make us expect him to be flexible in repeating a passage of his score, to interpret it differently or to sing his aubade not under a balcony but on a boat instead. In contrast, we may expect him to have great difficulty in changing particular notes in a fast melody at will. Apparently, we do assume that some action components are more stable and less flexible, then others. The study of errors or disorders that are visible in action performances can learn us more about our assumption or expectation – indeed, most errors in patients are considered to be extreme forms of normal errors that are made by healthy agents (Cooper 2002). What can we learn from errors and how do we normally avoid errors or unintended actions to occur?

Similar to the relative flexibility of motor representations that we have found in section III.2.2.1, relying on a template structure with open slots that permit variability, we should expect the representations of proximal intentions to allow some flexibility. Indeed, in a theory of action schemas preceding our CS/SAS theory the difference between closed and open skills was made. Whereas closed skills were taken to rely on a constant environment that permits identical performance of a particular action, open skills require a skilled performer to adapt his skill performance to environmental changes (Schmidt 1975). This and other types of variability have been integrated in Norman & Shallice's comprehensive network model of interactively activated components and are also responsible for action errors. Let us first look at the representations involved and how they would allow flexible adjustments, before focusing on the processes that potentially would carry these out.[396]

To begin with, different action hierarchies can share component representations,

[396] Lacking in the CS/SAS model are affective or motivational processes. In the –somewhat related - theory of Structured Event Complexes, an explanation is offered of agent's retrieval of action events from memory and how this impacts upon his behavior. It explicitly associates motivational and effective values with action components, explaining why actions usually appear to be satisyfing and why some patients would repeatedly engage in apparently aimless actions – like patients suffering from utilization or compulsive behavior (Huey, Zahn et al. 2008). The CS/SAS model could be expanded by letting such motivational processes also – indirectly, perhaps - modulate specific activation levels.

since a particular component action can contribute to more than just a single comprehensive action. Indeed, based upon this characteristic, so-called 'capture errors' can occur, when a particular action sequence shifts to another action due to the fact that they share a common action compnent (Cooper and Shallice 2000): boiled water is being poured in a tea pot instead of being used for making coffee, since both are components of the action representation of pouring boiled water. In this case, the activation levels pertaining to the irrelevant object (tea pot) and corresponding action goal (making tea) are apparently not decreased sufficiently to avoid this error from happening. For another, completing an action does at times require an alternative component that deviates from routine components to become flexibly integrated its performance (Cooper 2002): a flexible cook would not hesitate to improvise by using a tea pot when making coffee with no coffee pot present. Are such resources of variability of performing a complex action based upon the complexity of the interactive activation networks with their distinct roles for components like environmental conditions, higher-level action intentions, and lower-level motor intentions - random noise is still an important additional source.[397] Without such variability the choice between two equally activated, competing alternatives – using a left or right hand, for example – could lead to an impasse. Deciding how to specify the action would then require supervisory processing which would slow the action down and demand extra cognitive resources which are unnecessary if variability avoids such a stalemate. Random noise – or another source of variability, we could add - would ensure that activation levels of both hands are only rarely equal (Cooper and Shallice 2000).

With the representations allowing modifications via modulation of activation levels pertinent to component action representations, to environmental triggers and to action goals – or criteria for their fulfillment – the question remains which processes may be capable of reaching such results. As we've noticed above, there have been distinguished several supervisory processes that allow willed control of an action through overriding

---

[397] There is increasing evidence that random noise enhances the sensitivity and flexibility of processes in the brain. In contrast to systems characterized by linearity of their operations and interactions, so-called 'stochastic noise' can enhance the performance of the brain – operating largely as a non-linear system – as was shown in a random noise stimulation experiment with healthy subjects on a perceptual learning task, for example (Fertonani, Pirulli et al. 2011). More generally, it has been proposed that 'chaotic itineracy' occurring at several mechanism levels – from neuronal assemblies to social interactions – can explain how stability and variability appear to be interdependent (Kunihiko and Ichiro 2003). Similarly, Freeman has demonstrated how the complex dynamics of oscillations that obtain due to different neuronal firing patterns can lead to a signal to noise ratio that offers minor stimuli a limited chance to greatly affect brain processes (Freeman 2000). There are many sources of variability in the brain – indeed, the occurrence of stabilized patterns of synchronized neuronal activations should perhaps be more surprising than the occurrence of the opposite. Accordingly, it has been pointed out that most brain studies focus on task-related processes that explain some 5% of the brain's energy consumption, whereas spontaneous neuronal activity uses most of it. Fox & Raichle argue that instead of calling this variability 'noise', it to a large extent correlates with activity of the default mode network (Fox and Raichle 2007).

and modulation the process underlying routine actions. Supervisory processes can be recruited at the different phases or stages of an action performance, depending upon its temporal and situational conditions. A coarse distinction can be made between the preparation, implementation and assessment stages, each characterized by distinct processes contributing to the action (Shallice and Burgess 1996). For example, if time permits and the situation is familiar, an agent may set novel action goals or configure an appropriate alternative action representation - prepare himself for using another tool or omitting a particular component action - and thus modulate the representations involved in his action routines. Alternatively, he may prepare himself to note particularly well whether a specific component action goal has been reached, depending on which another component must not be performed (Shallice and Burgess 1996).[398]

According to the original CS/SAS model, such blocking of an action would require supervisory processes depending upon frontal lobe activity in constrast to automatic action (Norman and Shallice 1986). As a consequence, the limitations of frontal processes in terms of work load and speed would affect such willed actions, limitations that were also at stake in controlled versus automatic processing that we've discussed in the context of dual-process theories in chapter II.3 However, comparable to the discussion about the distinctions and relation between controlled and automatic processes, the relation between CS and SAS is under debate.[399] Research of action errors in in patients, for example, has demonstrated that some of their routine behavior is also affected by frontal damage, even though the original model suggested that such damage should only affect supervisory processes that take place in the frontal lobes. Apparently, it was concluded, routine actions do also involve some monitoring and correction, without necessarily requiring costly supervisory processes (Cooper 2002). A study of the effect of forming implementation intentions preceding the performance of stimulus-response actions confirmed that a supervisory process can preliminarily influence automatic action without slowing it down. A goal-intended action as a response to a novel stimulus was performed faster after forming such implementation

[398] From studies of the regulation of automatic behavior, it is known that priming can affect an agent's future perception of critical environmental features or his entertainment of certain goals, for example (Macrae and Johnston 1998).

[399] Research in subjects with high and low susceptibility to hypnosis suggests also that the supervisory processes in those with high susceptibility have weaker connections to CS processes. Subjects who are highly susceptible to hypnosis were more flexible in shifting their attention and responses in a Wisconsin Card Sorting Test, suggesting less influence of supervisory processes. In another test these subjects were also less capable of recalling details from the performed task (Aiktns and Ray 2001). These results confirm that supervisory processes can indeed impede flexible routine actions (Norman and Shallice 1986), but also that they may be required for establishing explicit representations.

intentions, even in frontal patients that suffer from utilization behavior or other disorders that tend to negatively affect response behavior (Lengfelder and Gollwitzer 2001). These results are taken to imply that automatic processing can be modulated at once following the reconfiguration of an intended action preceding a task.

In sum, studies of the interaction between automatic and controlled processing do indeed confirm that agents have the capability to block or modulate a proximate intention. However, they do also point out that multiple options for such modulation are available, each depending upon different processes and each targetting different components of the action representation and its underlying network. Corresponding to this, there are multiple constraints that limit the flexibility which is required from a responsible agent as he emerges from the philosophical analysis. This analysis has itself focused on some specific constraints of proximal intentions. Let us consider these as well.

3.2.3 Multiple processes and the constraints for anchoring an intention in a situation

In section III.3.1.3 we discussed constraints for anchoring an intention in a particular situation, which requires assessing whether that situation is appropriate for its execution and then specifying the action adequately. These constraints amounted to a demand for the action's internal consistency and for its specification in terms of "subplans concerning means, preliminary steps, and relatively specific course of action, subplans at least as extensive as I believe are now required to do what I plan" (Bratman 1987 31). Given the different processes involved in action control according to the CS/SAS model, the implementation of these constraints is manifold.

Concurring with the latter constraint, the model does ascribe to action representations a hierarchical structure which integrates many elements, among which environmental triggers, objects, component action representations, pre- and post-conditions, as well (Cooper and Shallice 2006). Once an action has been often practised, this structure and its elements are connected via interactive activation levels with each other. Depending upon the expert's experiences, his action routines will depend upon an elaborate representation with a great number of elements that can potentially be activated, allowing him to anchor the activated action representation rather flexibly in any given situation. A crucial feature of the model is that it also includes optional elements at several levels of specificity and of various kinds. For example, if the temporal order of an action sequence is flexible, the lateral activation and inhibition levels of neighbouring component action representations will be less decisive and thus allow the sequence to be influenced by the perception of relevant objects or the occupation of his hands, which can consequently – bottom-up - influence these activation levels (Cooper, Schwartz et al. 2005). When Don Giovanni and Leporello change clothes, the order of their dressing is dependent upon the availability of a cape or trousers, upon their having free hands or balancing upon a single foot, and so on: the anchoring and specification of the action is facilitated by the fact that the contention scheduling process can dynamically operate within a space of actions characterised by a complex structure of interactive activation levels that is sculpted through the accumulation of experience.

This explanation of how a proximal intention is carried out is quite different from the philosophical rendering of it. We've discussed how the philosophical account emphasizes that we should expect proximal intentions to operate between the domains of – implicit, experience based – motor intentions and distal intentions. Proximal intentions are then responsible for the process of specifying a distal intention such that it can be anchored in a given situation or, conversely, blocked in an exceptional situation. Without proximal intentions fulfilling these roles, it was argued, it is difficult to conceive how intentional actions can be performed. The implementation of these roles in terms of representations and cognitive and neural processes has shown that there is not a neat mapping of levels of intentions to the component processes suggested by neuroscientific research. For each intention type depends upon an explanatory mechanism of great complexity, containing distinguishable components which have not always recognizable counterparts in the results of philosophical analysis. Indeed, this lack of comprehensive correspondence between the empirical and philosophical accounts have to do not only with differences in aim and goal, but also with differences in the relevance attached to results of the other account.

Notwithstanding these differences, the CS/SAS theory and model can account for observations and simulations of routine actions that are in agreement with functional properties as presented in the philosophical analyses. A more direct similarity between the two perspectives is when explicit action representations are used by supervisory processes to control an action. Even though these processes are only capable of modulating routine actions, according to the model, this modulation can indeed lead to blocking an action wholesale, or altering targeted elements of it. Obviously, for such targeted modulations, the agent must indeed be experienced as he must be able to configure an alternative course of action, using a different object, following another sequence, and so on. A dissociation between such distinct supervisory processes has been proposed, together with different neural implementations for them (Shallice and Burgess 1996). Even though research has shown that imagery can also support such action preparations or modifications (Kosslyn 2008)[400], it is assumed that verbal representations will usually play an important role.

With regard to the former constraint, referring to the consistency of action, this would also be supported by the strength of the interactive activation patterns of an agent. Indeed, action disorganization or utilization behavior has been simulated by taking connections or elements of the comprehensive representation out or by adding random noise to the model, which amounts to blurring patterns of associated activation levels. As a result, the lateral inhibition of two inconsistent component action representations is decreased, for example, making it more likely that both of them are being performed.[401] These simulations have been compared with problems in neural connectivity – deviant neurotransmitter or receptivity levels, for example - and with the presence of specific neural lesions, responsible for the absence of elements in the relevant representations (Cooper and Shallice 2000).

With the later addition of pre- and post-conditions to the representation of each component action, the processes enhancing an action's consistency were further expanded. As these conditions can be considered to add to an action representation's activation level by taking into account activations based upon the presence of triggering (pre-)conditions and of inhibitory (post-)conditions, they further constrain its activation (Cooper 2002).[402] Given that these triggering situations and goal states are also relevant for maintaining consistency in routine action, the originally sharp distinction between supervised and automatic action processes has been weakened.

Indeed, research of frontal lobe patients demonstrated that different types of errors in routine behavior occur in correspondence to different lesioned areas, thus suggesting that routine action does also rely on some type of supervisory processing (Schwartz, Montgomery et al. 1998).[403] This finding did concur with another study

---

400 Kosslyn explicitly refers to mental imagery as a way to 'cognitively restructure' a stimulus or event, including its affective properties (Kosslyn 2008).

401 Utilization behavior has indeed been associated with a lack of inhibition of impulsive action due to lesion of the frontal lobes (Archibald, Mateer et al. 2001). Indeed, the author who coined the term 'utilization behavior' emphasizes how subjects with this disorder are 'abnormally dependent' upon stimuli from the social and physical environment (Lhermitte, Pillon et al. 1986). The lack of inhibition and increased sensitivity to external stimuli together constitute the disorder's most important symptoms.

402 Presumably, the perception and recognition of relevant environmental triggers and the recognition of having reached an action's goal state depend upon activations of distinct neural networks. Mirror neurons can play different roles in these processes. Especially for object related actions, mirror neuron networks have been found that are sequentially activated during different phases of such actions – from the trigger phase of an action to its goal state. For that reason, some mirror neurons are said to be 'logically related' (Iacoboni, Molnar-Szakacs et al. 2005).

403 Denying the sharp distinction between two processes responsible for automatic and willed control of action, Schwartz et al. explain the action errors in automatic control via a lack of cognitive resources (Schwartz, Montgomery et al. 1998). Such a lack perhaps obtains more often in automatic control, as distractor objects may play a greater role under such conditions than in willed control, for example.

of differentially impeded action representation capabilities in patients. Errors in the organization of an action script were found to be associated with frontal lobe lesions, whereas the generation of a new action script appears to rely on semantic networks that are located in more posterior areas (Sirigu, Zalla et al. 1995). Apparently, in both types of processing, action consistency is supported by several cognitive processes that each contribute to different forms of action consistency like its appropriate initiation, its completeness, its correct order and its adequate completion.[404]

In sum, the adequate activation of a comprehensive, hierarchical representation of a complex action together with the appropriate activation levels that connect action components, triggers and goal states with each other, can be impeded in multiple ways as it is dependent upon a rather large network of neural areas interacting with each other. Flaws in these processes are observable in an agent's behavior, for example in his failure to perform a sufficiently specified action or to do so in a consistent way. As noted before, there are strategies an agent can employ to prevent this, like when he forms implementation intentions (Gollwitzer and Sheeran 2006), engages in counterfactual thinking (Galinsky and Moskowitz 2000) or employs mental imagery (Kosslyn 2008). However, such strategies require that an agent is capable of a preliminary specification of relevant action elements. For only on that basis is it possible for an agent to modify or reconfigure such an action. This reminds us of our extensive argument in Part II, that although expertise corresponds with kludge formation and some modification of the representations in use, it cannot imply the loss of the capability to articulate the representation that is used in the expert's skill. Without the latter, an expert would be less capable of modifying, adjusting or correcting his actions than a novice, which appears to contradict the notion of expertise itself.

Let us consider in the final section devoted to proximal intentions some neural evidence pertaining to them and see whether we find in this context kludge formation to play a role as well, affecting the mechanisms responsible for them.

3.2.4 Proximal intentions and some evidence concerning their neural implementation

The motor representations and the neural processes that are involved in motor

---

[404] Research in comprehension of texts during reading has motivated the development of an interesting 'landscape' model, in which interactive activations – here of concepts - play an important role as well. The authors demonstrate both in empirical and simulation studies that readers employ two mechanisms that together constrain incorrect or incoherent interpretation: associated concepts are being activated while a parallel coherence-based retrieval process can help to detect errors (Tzeng, Broek et al. 2005 ; van den Broek and Kendeou 2008). Given that processing of language and action overlap to a large extent with each other, it would have been interesting to further compare this 'landscape' model with our model of a 'sculpted space'.

intentions were found to undergo changes as a result of gaining expertise, which leads to what we've been calling 'kludge formation'. Earlier in this chapter we've described how a phase of increasing efficiency in neural activations is followed by a phase in which representations become increasingly associated with other representations or processes that allow a wider range of exploitation of the motor representations. The representations itself do develop into more complex ones, as well, and develop a template structure in which information is compressed in the form of chunks in combination with open slots to allow for the integration of variable information. With proximal intentions and their involvement in anchoring and specifying an action in a given situation and in blocking it in exceptional cases and two different computational processes responsible for them, it is plausible to expect that their neural implementation is more complex, too.

Norman and Shallice's theory of willed and automatic control of action has also been discussed with regard to its plausible neural implementation. Obviously, Hebbian learning and other processes that affect connectivity are put central, since the theory posits that every action depends upon a specific configuration of activation and inhibition levels of a hierarchically structured set of action representations, eventually leading to the selection and execution of a particular action (Norman and Shallice 1986).[405] In the theory's expanded version we learn that after a period of skill learning a hierarchical action representation has been established and that this complex action is also assumed to be somehow represented neurally, for "even in a domain as loose as the organization of everyday routine action, one cannot simply dispense with *units* or *discrete states* representing action subroutines and goals (Cooper and Shallice 2006 906, italics added). However, the specification of this neural implementation of an action representation is here left open.[406] At the same time, the theory also predicts the emergence of such representing units or states while equipping these representations with various kinds of open slots. For it is with these open slots that an expert action – characterized by stable and efficient activation patterns - can flexibly vary in the sequence of its component actions or can respond to several potential triggers or make

[405] Indeed, it is suggested that when the CS/SAS model is used for simulation studies: "a parameter may be taken to correlate with the level of a neurotransmitter", or a "parameter may be related to the ratio of the connectivity from the two activation pathways to schema nodes" (Cooper and Shallice 2000 323).

[406] This remark of the necessity of the representation of action subroutines and goals by units or discrete states is aimed against an alternative computational model that aims to avoid a representation of a action's hierarchical structure. It is implemented in a recurrent connectionist network to simulate action sequences and disorders of action, making it a plausible alterantive according to (Botvinick and Plaut 2004). It is debatable, however, whether this alternative model can equally account for the supervisory processes or other strategies for action modulation that rely on explicit and specified action representations (Cooper and Shallice 2006 906). We've discussed a similar debate with regard to learning and development and the role of – both implicit and explicit - representations in Part II.

use of different effectors (Cooper and Shallice 2006). As a result from this first take on their neural implementation, it appears that proximal intentions – mediating between motor and distal intentions – are also characterized by the formation of kludges and the development of representations with a template structure. Let us consider more specifically how these requirements could be neurally implemented.

Neural implementation of the original theory was thought to require two distinct neural mechanisms, one for automatic action control and another for the supervisory processes involved in willed control (Norman and Shallice 1986). This distinction was challenged by studies which showed that frontal patients were not only impeded in such willed control but also in their routine actions, suggesting that the two mechanisms were less distinct than previously assumed (Schwartz, Montgomery et al. 1998). Such pathological evidence in combination with animal, developmental, and neuroscientific evidence has led to the proposal that prefrontal cortical activations are generally involved in developing, storing and processing hierarchically structured representations that are also being employed during different types of action processing. These representations are called 'structured event complexes' or SECs and they can vary in several dimensions (Grafman, Sirigu et al. 1993).[407] As this SEC theory challenges the CS/SAS theory particularly with regard to its neural implementation and not so much with regard to its assumptions regarding the representations involved, let us pause here for a minute with it.[408]

The SEC theory does concur with the CS/SAS theory in that Hebbian learning is vital for the development of the representations, depending as they do on the frequency of co-activation of features or items, or their similarity or association value (Grafman 1995).[409] The SEC theory, however, further elaborates the representations and puts them more central in a larger group of cognitive processes, even though these representations rely so much upon prefrontal cortical activations which the original

---

[407] An 'event' in this context is defined by its being thematically consistent although it still can range from a simple motor movement to a more complex event. The SEC is 'structured', as it contains several components in an orderly fashion and with temporal constraints and it is 'complex' because of its consisting of several components that are assembled (Grafman, Sirigu et al. 1993). Here, again, a gradient of complexity is present with less complex SECs being stored in posterior areas in contrast to more complex SECs in anterior areas (Forbes and Grafman 2010). The structure here is understood to be hierarchical, too.

[408] The initial version of the SEC approach did make similar distinctions as the later ones. However, in that version SECs were considered to be only at the bottom of the hierarchy of 'Managerial Knowledge Units'. SECs were then considered to be the 'developmental and phylogenetic precursors' of these MKUs (Grafman 1995). Meanwhile, MKUs do no longer figure in the theory and SECs are considered to become more complex as a result of developmental maturation and especially PFC development.

[409] Indeed, frequent co-activations do generally lead to sparser, simpler representations which in turn can affect subsequent decisions and actions in terms of processing speed and stability (Grafman and Krueger 2008). This concurs with what we've discussed in the previous Part regarding the process or representational redescription as part of learning and developmental processes.

theory considered to be primarily responsible for supervisory processes only.[410] For example, the CS/SAS theory does offer both in its original (Norman and Shallice 1986) and in a later version (Shallice 2002) room for the spontaneous generation of schemas, supported by PFC activations, but limits this capability to the modulation of those lower-level schemas that are employed in contention scheduling. Compared to that, the SEC theory suggests a still larger human capability of establishing event representations at will: "[t]he SEC could be established on the basis of experiencing external events or through the generation of "internal thought"" (Grafman 1995 348). We can expect to see that this human capability of spontaneous schema generation or internal thought and the consequences of such schemas on subsequent behavior is important for our discussion of distal intentions.

Important to note here is that growing expertise – relying on both automatic and controlled processes – affects a range of related processes, including the perception and recognition of novel events. This is demonstrated convincingly in studies of a specific task with regard to event processing, namely event segmentation. Expertise amounts to memory storage of an increasing amount of ever more complex representations, which are modulating subsequent perceptual processes. As a result, increasing expertise corresponds with faster and more accurate recognition of the structure of an action, including its fine and coarse segment boundaries and the hierarchical relations between segments (Kurby and Zacks 2008).[411] Event segmentation research concurs with the SEC theory in that observers primarily recruit prefrontal cortex activation during segmentation tasks, as these activations increase significantly when observers not just passively observe events but have to judge them. However, this research suggests additional recruitment of parietal cortex, specifically for representing temporal features of the events (Zacks, Speer et al. 2007).[412]

---

410 The contribution of Grafman and others to explanations and theories about action processing has been taken as a further elaboration of the representations involved in CS and SAS processes. Particularly the addition of information about temporal duration and about relative importance to (component) action representation has been recognized as a welcome addition (Cooper 2002).

411 Event segmentation research is inspired by research of reading comprehension that relies upon readers' construction of situation models in which events and intentional actions play primary roles (Zwaan, Langston et al. 1995). Corresponding to theoretical developments that can be observed in other domains of research of cognition, Zwaan et al. have claimed that the representations found to be involved in reading are in fact employed for a much wider range of mental simulation tasks (Zwaan 2009).

412 Event Segmentation Theory claims that event representations play an important role in the predictive coding that the brain is constantly doing. Moreover, when errors are detected between predictions and actual observations, the representations are allegedly updated. Apart from its articulation of some additional uses of the representations, the EST has also been presented with an elaborate neural implementation, accounting for many features of event segmentation processes and the features that matter in these. As a result, several cognitive disorders – ranging from Alzheimer's dementia to obsessive compulsive disorder – are partially explained according to this framework (Zacks and Sargent 2010).

Nonetheless, event representations recruit primarily prefrontal areas and these play a central role in the SEC theory, according to which these representations also represent more information than those included in the original CS/SAS theory. This is enabled by the fact that neural implementation of SECs in prefrontal cortex implies their having richer connections to other neural areas. For example, these representations generally contain elements with a modality specific origin.[413] Notwithstanding such an origin, SEC components can still be activated via internal generation or simulation once they have become associated with other and more abstract components. As a result, such an action simulation will still contain properties that are due to its modality specific origin (Grafman and Krueger 2008).[414]

Apart from the modal features that are included in event representations, SECs representations are also considered to have social and emotional features, as these are observed to play a role in action selection processes as well.[415] This again expands their distributed neural implementation. SECs are therefore also associated with activations in cortical or subcortical structures, enabled again by the rich connectivity that neurons from ventral, ventromedial and medial prefrontal cortices have (Grafman 2006 ; Krueger, Barbey et al. 2009).[416] Indeed it is not only the rich connectivity

---

413 Ridderinkhof et al. present in their review of neurocognitive processes involved in control also an overview of several cognitive processes that are involved in control, like processes aimed at goal-states, anticipation of reward, performance monitoring, error-detection. They also contend that the PFC plays a major role in all these components of control, facilitated by its high connectivity to other neural regions (Ridderinkhof, van den Wildenberg et al. 2004).

414 In this respect, the SEC theory concurs with accounts of embodied cognition like Barsalou's account of simulators that we discussed in section II.4.1.1 For example, SECs can be considered as representations that underly simulations of a more specific and restricted kind, given that they represent primarily events. Indeed, Grafman and others elsewhere develop 'elators', that is 'event simulators', while referring to Barsalou's work (Krueger, Barbey et al. 2009). However, Barsalou's account explicitly describes how component representations are also stored in sensori-motor areas and not solely in PFC (Barsalou 1999c), which is distinct from the SEC theory.

415 Reward values have been associated with SECs as well. The explanation of an obsessive compulsive disorder would accordingly be that, due to a neural pathology, an agent may not experience the reward associated with the completion of a hand washing SEC and thus feel the urge to repeat the action corresponding to the SEC (Huey, Zahn et al. 2008). This can be considered an extension of the notion of a component action's 'post-condition' as presented in (Cooper 2002), which was there not associated with reward value experience.

416 In his account of moral cognition, Moll and others have integrated SEC representations with such emotional and motivational affects, acknowledging the importance of the latter for making moral decisions (Moll, Zahn et al. 2005). Obviously, PFC components do play a role in moral cognition, as is evident from ventromedial PFC activation being involved in social stereotyping, for example (Milne and Grafman 2001). However, other areas also play a role as for example the temporo-parietal junction, which is involved in choosing between selfish and altruistic acts (Morishima, Schunk et al. 2012). Generally, acting according to the results of social (or moral) cognition, though, appears to be dependent upon further motivational contributions. Based upon this, Frijda emphasized the role of emotions for an agent's 'action readiness' (Frijda 1986).

of prefrontal neurons, but also their capability of sustained firing and the fact that prefrontal pyramidal cells are more spinous than those in other areas that supports SEC characteristics (Grafman and Krueger 2008).

With a predominant role for prefrontal cortex according to the SEC theory, it does not present us with a discrete distinction between controlled and automatic processing like the CS/SAS account. It still acknowledges a gradual process of automatization, corresponding allegedly to both the strength of the event representation and its increasingly sparser, economic coding. This process corresponds neurally to a shift from anterior to posterior PFC and facilitates the execution of the representation via neighbouring motor areas (Grafman and Krueger 2008). With such reliance on PFC as areas for storage of SECs, subcortical areas like the basal ganglia are only implied when such an execution takes place (Grafman, Sirigu et al. 1993).[417]

The kludge formation that we associated with expertise is a relatively circumscribed process according to the SEC theory: stronger representations are coded more economically and in more posterior regions (Grafman and Krueger 2008).[418] The CS/SAS theory is in agreement with the relevance of changing activation levels for expertise, with associated action representations becoming ever more complex yet requiring less neural activations.[419] Indeed, both theories agree that as a result, expert action and cognition can become more complex and flexible, yet also be faster than in a novice. In contrast to the SEC theory, the original CS/SAS theory suggested that automatic action does not equally rely upon prefrontal cortical activations.[420] Although the strict separation between CS and SAS processing may have been abandoned, the theory does still make a distinction between the two processes with the process

---

[417] Obsessive-compulsive disorder is explained accordingly with the basal ganglia setting too low a treshold for the activation of prefrontally stored SECs, thus causing their undesirable motor performance (Huey, Zahn et al. 2008).

[418] Apart from this more general prediction based upon the SEC theory, studies of planning and script analysis with patients has suggested that in fact several distinguishable processes like sequence ordering, categorization, and script generation rely upon different prefrontal areas (Sirigu, Zalla et al. 1996 ; Sirigu, Zalla et al. 1995). Indeed, Grafman even suggests a lateralization for integration across events (right PFC) and single event integration (Grafman and Krueger 2008).

[419] Surprisingly, Grafman purports that the CS/SAS model is primarily a processing theory and less interested in the representations involved (Grafman 2006). In light of the extensive computational studies, aimed at simulating empirical evidence of normal and disordered processing of a specific action (preparing coffee) (Cooper 2002 ; Cooper and Shallice 2000), this critique is misdirected. Indeed, these later versions of the CS/SAS model are more explicit with regard to the processing of representations than Grafman's SEC theory is. As the latter is more detailed with regard to possible neural implementations, the integration of both theories promises interesting results.

[420] As mentioned above, an early critique of the assumed separation of automatic and willed control came from patients with frontal lobe lesions which had difficulty in the ordering and the generation of action scripts – the former task depending upon more anterior areas and the latter on semantic networks that are located in more posterior areas (Sirigu, Zalla et al. 1995).

responsible for automatic action recruiting less frontal areas.[421] To end this chapter on proximal intentions, we will consider neural areas that may underpin the automatic process of action determination.

The process of contention scheduling, underlying automatic action control, was originally suggested to be distinct from controlled action, and to rely completely upon "mechanisms in the corpus striatum of the basal ganglia, often thought to be involved in the selection of actions" (Norman and Shallice 1986 10). Formulated concisely, these mechanisms are responsible for the activation or inhibition of representations that support or conflict with a particular action. In healthy subjects, upon the – willed or automatic – activation of a high-level action representation, a range of related component action representations ought to be activated and inhibited. Later, some specification of these mechanisms was provided.

One mechanism that was mentioned is the striatal dopamine system, which is involved in such activations in the CS processes and of which the deficiency – in Parkinson's disease, for example – is observable in slower initiation of an action. Another mechanism is at work when deficiency of the amphetamine system is associated with failures in inhibiting (component) action representations, observable as the simultaneous performance of multiple, conflicting actions or the repetition of an action that has just been completed (Cooper and Shallice 2000). These disorders are determined by mechanisms that are influencing activation levels in a rather indiscriminate way and thus disturbing the more specific modulation of activation levels of those neural networks that are associated with a particular action representation – whether located in PFC or not.

Cooper later contended that in healthy subjects, it may also well be that the basal ganglia are involved in modulatory processes that lead to selection and inhibition of respective representations even in non-automatic control (Cooper, Schwartz et al. 2005). This contention was partly motivated by evidence that basal ganglia are indeed playing generally an important role in the allocation of neural resources for

---

421 Research with parietal cortex patients showed that these suffer from problems with specific action script generation tasks, suggesting that relevant representations do indeed not only rely upon frontal areas – even not in controlled processes (Godbout, Cloutier et al. 2004). Other research, particularly aiming at specification of the functions of mirror neuron systems, suggests that representations of relatively simple motor actions for their performance and recognition are stored in premotor cortex (PMC). Patients with ideational apraxia appear to have difficulty in activating the CS scheduling system such that these motor intentions stored in PMC are activated for the recognition of motor actions. It confirms the involvement of CS also in recognition processes and the involvement of regions outside PFC (Rumiati, Zanini et al. 2001). Other research with frontal patients did show that their action representations were not affected like the SEC theory would predict, as it locates these representations primarily in PFC. The effects of the frontal lesions in the patients' performance of verbal and pictorial script tasks were taken to be the consequence of a decreased capability of rejecting wrong alternative responses, which relies upon PFC. This was taken to support the CS/SAS theory and not the SEC theory (Zanini, Rumiati et al. 2002).

the performance of cognitive and motor tasks (Redgrave, Prescott et al. 1999).[422] Even though the strict separation of neural correlates of CS and SAS processes has been rejected and some overlap between the two has been acknowledged, it is argued that in light of evidence – including evidence from patients and from the design of autonomous agents – this account is still presented as a dual-process theory (Cooper 2002).[423]

In closing, it may be useful to return to a model of action selection processes that we referred to in section III.2.2.1. It was discussed there in the context of our discussion of motor intentions and emphasized how multiple motor intentions could arise upon the perception of environmental affordances, requiring the agent to – implicitly, or not – select or decide between those competing action options (Cisek and Kalaska 2010). Indeed, we used the model to confirm our argument that, based upon experience, a form of kludge formation contributes to influencing the action affordance competition such that in experts a sculpted space of actions can be observed. Now when reconsidering this model for sensorimotor control once more, we recognize how it ascribes a crucial role to the interactive activation of multiple, hierarchically structured action representations, as well. The question is, whether the model could be expanded to account also for proximal intentions and in such a way that it integrates some form of secondary or supervisory processing. In that case, it would look similar to the structure of the CS/SAS model.

Indeed, this account of motor intention processing does leave explicitly room for supervisory processes that affect or modulate the affordance competition process which it describes. Indeed, it does so in line with our earlier reasoning in Part II that learning and development tends to employ representations and processes that have already been established earlier, which as a result become even more deeply entrenched. On

---

[422] Based upon interdisciplinary lines of evidence, it is proposed that the basal ganglia have evolved and are particularly well suited for selection problems in general. Moreover, they operate at several levels of specificity of the cognitive or motor options that compete for selection, which is confirmed by the association between their dysfunction and particular disorders in cognition and action (Redgrave, Prescott et al. 1999). Another review equally assigns an important role to basal ganglia activation, in connection with cortical (pre-motor) activations for the generation of relevant action representations (Graybiel 2008). Conversely, for the inhibition of an automatic response an agent must be capable of modulating basal ganglia activations (Aron 2011).

[423] The effect of action familiarity has been studied from other theoretical perspectives, as well. It has been found that familiar action representations are not only benefitting action responses but are also shared with cognitive or semantic tasks. Although not explicitly using a dual-process approach, it is acknowledged that high-level planning does have access to such shared representations as they are capable of activating them, even though the planning activities themselves are both cognitively and neurally distinct (Elk, Viswanathan et al. 2012). Such studies build partly upon the theory of event coding that denies a strict separation between codes that are used in perception and in cognition or action (Hommel, Musseler et al. 2001). Needless to say, that the CS/SAS model is concurring with the notion of shared representations but does emphasize how these may be accessed and employed differently during different task performances.

the basis of their model the authors suggests that other action selection processes are in fact modulating the hierarchically structured selection processes as explained by the model: "[t]he recent evolution of primates is distinguished by advances in the ability to select actions based on increasingly abstract and arbitrary criteria" (Cisek and Kalaska 2010 283).[424] Such controlled selection occurs by way of biasing or strenghtening a particular option, which takes more time than the automatic selection of an action option and does indeed rely primarily upon many forms of PFC activations.

In sum, our discussion of empirical evidence concerning proximal intentions did have a result that concurs in broad terms with the previous philosophical analysis. The latter did suggest that proximal intentions are in fact mediating between motor intentions and distal intentions and in doing so responsible for the anchoring or specification – or blocking, for that matter – of a distal intention in a given situation. Indeed, the distinction of separate proximal intentions has been questioned as such since action determination could allegedly be analyzed in terms of motor and distal intentions alone (Holton 1999). The cognitive neuroscientific model of willed and automatic action which was more closely looked at in these sections concurs largely with that analysis, dividing the task of action control between two distinct types of processing.

If expertise does indeed affect an agent's sculpted space of actions, it must therefore occur through these processes and their interaction. The present chapter has presented insights that proximal intentions are also affected by the modifications that motor intentions undergo as a result of learning and development. To the extent that proximal intentions are to determine action, they can do so mainly by modulating the interactive activations that in a given situation spread throughout a hierarchical network of stabilized – redescribed – motor action representations. The supervisory processes that create these modulations are indeed able to sustain and enhance a coherence that transcends the aims and criteria of the particular situation.

With that, we already touch upon the distal intentions. In the philosophical model presented by Bratman and adopted in modified form by Pacherie, these distal or 'future-directed' intentions are given methodological priority. The implementation of such a priority would amount to a top-down determination of action, undisturbed by bottom-up processes like those which are responsible for motor intentions. In reality,

[424] One criterium that is important is of course the expectation of reward. However, there are often several rules applicable to a single situation. Dorsolateral prefrontal cortex may play an important role in favoring one over another rule, with correspondingly different action preferences. This is just one of many potentially relevant criteria, with PFC playing a crucial role in their implementation (Ridderinkhof, van den Wildenberg et al. 2004).

we have seen already that top-down control appears to be limited to modulations of activations of action representations at lower levels. How, then, can distal intentions contribute to long term processes, to an expert's sculpted space of actions? Answering this pertinent question, we will in the next and final chapter of Part III' focus on the issue how an agent can sculpt his space of actions by the controlled (re-)configuration of actions via the explicit articulation of distal intentions. Through such articulation he is both equipping his supervisory processes with rules and criteria that are relevant for proximal intentions and simultaneously modulating the activation levels of the motor intentions that need eventually to be executed.