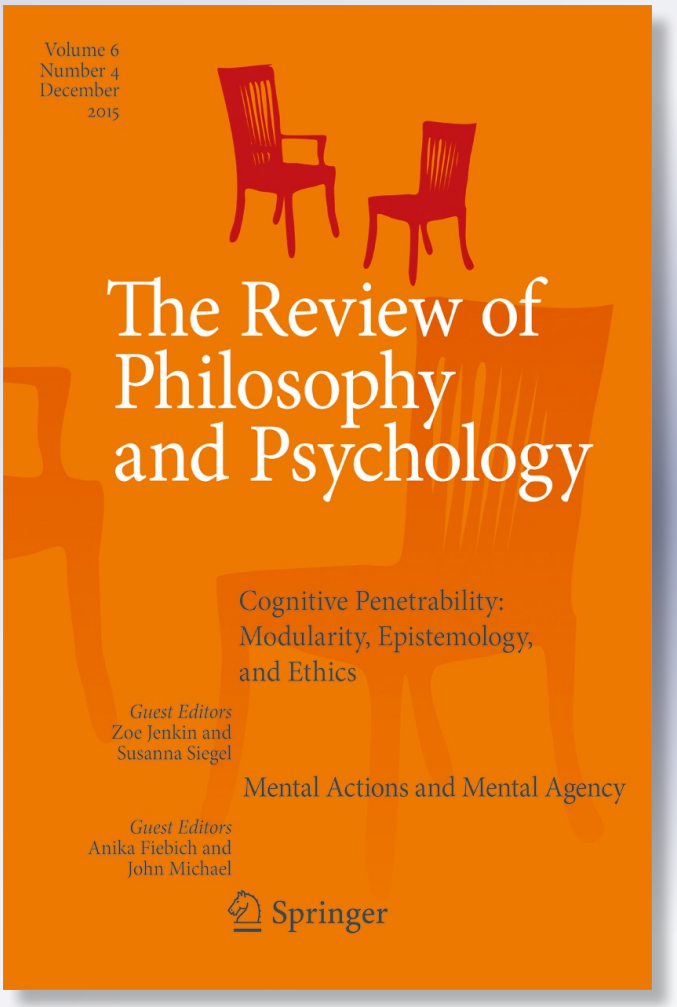
Time and Action: Impulsivity, Habit, Strategy

Joëlle Proust

Review of Philosophy and Psychology

ISSN 1878-5158
Volume 6
Number 4

Rev.Phil.Psych. (2015) 6:717-743
DOI 10.1007/s13164-014-0224-1



Springer

Rev.Phil.Psych. (2015) 6:717–743
DOI 10.1007/s13164-014-0224-1

Time and Action: Impulsivity, Habit, Strategy

Joëlle Proust

Published online: 15 November 2014

Abstract Granting that various mental events might form the antecedents of an action, what is *the* mental event that is the *proximate* cause of action? The present article reconsiders the methodology for addressing this question: Intention and its varieties cannot be properly analyzed if one ignores the evolutionary constraints that have shaped action itself, such as the trade-off between efficient timing and resources available, for a given stake. On the present proposal, three types of action, impulsive, routine and strategic, are designed to satisfy the trade-off above when achieving goals of each type. While actions of the first two types depend on non-conceptual appraisals of a given intensity and valence, strategic intentions have a propositional format and guide action within longer-term executive frameworks involving prospective memory.

1 The Fundamental Problem in the Philosophical Action Theory Seen from an Outside-in Viewpoint

Most naturalist philosophers of action have tried to account for action in causal terms, and recognized that flexibility and adaptiveness to context suggest that goal representations must play a causal role in selecting and guiding behavior. Belief-desire theories first seemed to score higher in accounting for the causal influence of attitudes on action, and, in particular, for the major role of practical reasoning in deciding to act. As it has appeared progressively, however, beliefs and desires are 1) neither necessary to trigger an action (as in emotional, arational, routine, or compulsive actions), 2) nor sufficient (as in actions having a so-called deviant causation, or in certain forms of the autistic syndrome and other pathologies of the will). Furthermore, framing one's goal in purely propositional terms has been shown 3) to fail to explain how a given intention to act is supposed to be executed, and 4) fall short of providing an account of action failure.[1]

Based on the assumption that various mental events might form the causal antecedents of an action, then, the question was raised of what is *the* mental event that is the *proximate* cause of action, i.e. both necessary and sufficient, able to guide behavior until termination, and to explain action failure as well as action success? Having raised

[1]For a review of the objections raised against belief-desire causal theories of action, see Pacherie (2002).

J. Proust (✉)
Institut Jean-Nicod (DEC, ENS), Paris, France
e-mail: jproust@ehess.fr

this question, which he called "The fundamental problem in Philosophical Action Theory", Sellars (1973) proposed that the proximate cause of an action consists in "intending to do something here and now". On his view, intending is a genuine propositional attitude with a world-to-mind direction of fit (in Searle's terms[2]), whose function is to mediate reasons and actions. While the debate raged for some time over whether intentions were bona fide mental states or mere stand-ins for beliefs and desires, they were finally found not only to have better explanatory adequacy than the latter, but also to be more easily compatible with the psychology and the neurosciences of action.

While recognizing the importance of Sellars' seminal paper and of the various subsequent refinements of the theory of intention,[3] these studies used a debatable method: they have concentrated on the representational function of intentions, rather than on the function of acting. The analysis of function, then, was conducted from inside out, from mental state to behavioral and distal outcome, rather than from outside in, i.e., from desired outcome to type of intention. Granting that our mental states have been selected to serve survival imperatives, we should ask not: "What is the representational content of intentions?" but rather: "What is the function of actions that explains and rationalizes the representational structure of intentions?"

Taking this outside-in viewpoint leads to an interesting reverse-engineering hypothesis. In the agent's attempt to change the world through goal-directed behavior, an action is subject to a trade-off between resource consumption and expected beneficial effects. Acting, from this viewpoint, can be considered as a homeostatic process in a dynamic system. Its function, then, is to regain an antecedent balance, lost through or threatened by energetic depletion, environmental changes, etc. The specific depletion it responds to constitutes a need, ranging from primary drives for food, shelter, mates, to more elaborate aims such as acquiring or extending one's social status, trading goods, learning competences, securing future gains, and so on. Resources available to restore homeostasis can be internal or external, i.e. under endogenous[4] or external control.[5] They will objectively vary across contexts and over time, as will the needs and their urgency.

In virtue of the cost-benefit trade-off, an organism should only be motivated to act when the resources used in order to attain its goal are not likely to exceed, all things considered, the gain likely to result from acting (considering both the risk incurred and the benefit expected). For example, is it worth trying to get food or water when there is a high probability of being shot when crossing the street? The final decision depends both on the intensity of the need and the assessment of the probability of the risk. There is another structural aspect of this trade-off, however, related to the timing of action, that has received little attention from philosophers. *No single agentive mechanism is able to respond simultaneously to (i) fast-moving objects, to predators, to sudden changes* etc.*, (ii) to recurrent events, and (iii) to long-term anticipated changes and needs*. A plausible hypothesis, which will be articulated below, is that this difference has generated three different action systems, with their specific representations and

[2] See Searle (1983).
[3] See Bratman (1987), Mele (1992), Pacherie (2006), Proust (2003), Searle (1983).
[4] E.g. attention available, level of motivation, presence of acquired know-how.
[5] E.g., instrumental mediations, cooperation with others, temporal constraints.

incentives, informational processes, and dispositions to act. These three systems, on this hypothesis, compete to control behavior as a function of the temporal properties of the context. The three fundamental time requirements have respectively to do with i) having here and now to respond to an urgent, unpredicted environmental change, ii) having here and now to take advantage of a recurrent pattern of opportunities, and iii) having to plan strategies to bring about changes later in time. These three temporal patterns correspond respectively to what we will distinguish as *impulsive*, *routine* and *strategic actions*.

Proceeding outside in, we will aim to show how each type of action is triggered by an intention, whose representational format and causal influence depends on the distal action to be performed. We will explore whether this tri-partition similarly applies to physical actions (those having as their goal to change the external environment) and to mental actions (those having as their goal to change the informational or cognitive dispositions of the agent). We will also investigate how a mode of action is selected given a context. We will, in particular, examine how the three types of action can be conjoined at different steps into a complex action, in order to meet conflicting temporal and goal constraints.

2 Types of Physical Action

Let us first examine how our primary trade-off might have shaped types of world-directed, physical actions. How, in other terms, did different temporal constraints associated with a disposition lead to the selection of a specific form of triggering event, having a particular representational format?

2.1 Impulsive Action

Examples of impulsive actions are running away out of fright, shouting when watching a game, ducking to avoid a projectile, all the other things one does without having planned to do them, sometimes without even knowing that one is doing them it until they are done. Calling them "impulsive", rather than "emotional" has the merit of stressing their relation to urgency. Impulsivity describes what Nico Frijda (1986) calls "control precedence": an agent intentionally performs her impulsive action here and now, possibly interrupting another ongoing action, and with no apparent concern for the consequences. Impulsive actions, then, are "reflex-like" and shortsighted: they react to some particular features of a present situation, ignoring others. A flat tire, a broken window: the angry owner utters a swear word, ignoring his intolerant hearer. Most authors have suggested that these actions are too quickly formed to respond to beliefs and associated desires.[6] They are rather done "in the grip of an emotion". For this reason, they have been claimed to be "arational" rather than irrational: although intentional, the agent did not do it for a reason. This does not mean, however, that impulsive actions are not adaptive. What, then, is their function?

[6] See Hursthouse (1991), Pacherie (2002), Scarantino (2014), Searle (1983).

2.1.1 The Function of Impulsive Action

The promptness of the disposition to react that they embody must be part of this function. We propose, then, that the selected-for function of impulsive reactions is to quickly trigger an appropriate response to an unexpected, urgent threat or opportunity. Following Frijda (1986), Andrea Scarantino similarly proposes that the primary function of an impulsive action is to prioritize an action tendency, to a degree, as a function of the kind and intensity of the associated emotion. In his proposal, however, flexibility seems hardly compatible with urgency: his claim is that emotions have "abstract relational goals", that need to be situated in a concrete context in order to guide bodily changes. An impulsive action, then, inherits its function from the function of its precipitating emotion (as derived from its evolutionary etiology). Fear carries information about predator because such is the function of fear.

The phenomenology of an impulsive action may be used to refine this teleological approach. An agent's attention and action tendency are both strongly focused *on* something—*this* threatening property, at a specific location, and on the ways of reacting appropriately. This requires a form of thinking, immediately and flexibly shaped both by the characteristics of the environment, and by the gestural repertoire of the agent. The abstract relational core of an emotion is a step in a theory, not a psychological episode in the agent's mind. These observations suggest that the function of impulsive actions can be read from the informational content that is used, here and now, to trigger them.

2.1.2 The Informational Content of Impulsive Representations to Act

According to a claim defended at length elsewhere, an impulsive action is set off by a feeling, i.e., a mental episode extending from perceptual uptake and appraisal to action tendency.[7] "Feeling" denotes a reactive, subjective experience with a distinctive embodied phenomenal quality. "Reactive" means that feelings consist in appraisals of a present property or event—called here "an affordance"—with their congruent disposition to act. Affordances are sensed in a non-doxastic way. In other terms, feelings are affordance-sensings: they track the relations between the agent and the properties in the environment that are meaningful to her here and now. The type of experience that is entertained when detecting an affordance has a given valence (positive, when discovering food or water, or negative, when detecting a predator) and intensity: these are two essential parameters of non-doxastic appraisals that determine how to react to the affordance.

On the present view, the feelings that lead to an impulsive action are of an affective, emotional kind (we will see later that non-emotional affordance-sensings generate non-impulsive types of reacting). Why, then, is this structure appropriate to the urgency of a threat, or of an opportunity to be seized? An "affordance sensing", like any appraisal, is constructed around the outcome: it includes in its representational content a motivation to react here and now, as a function of the spatial information registered. Its being non-conceptual allows the appraisal to be made quickly and in a modular way (merely from its input). Affordances obviously constitute only a part of the information that the agent may use. As our temporal trade-off has suggested, however, the information they

[7] For a detailed defense of this proposal, see Proust (2014a).

convey is cheap to collect, quick to process, and, potentially of major significance for survival.

The represented affordance includes, then, all the main elements relevant to acting, and, by default, will spark an impulsive action that corresponds to its valence and intensity (if the intensity is low, the disposition to act may not result in an action). The representational structure of a feeling thus includes input, time, location, and a situated disposition to act:

(1) Feeling-representational content: Affordance_{a} [Place=here], [Time=Now/soon], $[\text{Valence}_{a=+}]$, $[\text{Intensity}_{a=.8}$ (comparatively specified on a scale 0 to 1)], [immediate disposition to act of degree_{d} with action program_{a}].

The temporal features of this representational system need to be emphasized. Affordances are typically detected—as a piece of self-relevant or salient information—on the basis of predictive cues, rather than of full-blown perceptual identification. Affordance predictions are made only milliseconds after visual sensations register on the retina, i.e. before the categorisation of perceived objects is completed.[8] If an agent had to wait for a conceptual identification of an object moving toward her head, she would not have the time to duck. As fencing trainers know well, (a sport where speed and timing play a major role), trainees must blindly practice, again and again, a particular parry until it is automatically, i.e. "thoughtlessly", associated with a sensed affordance. Uptakes of affordances, in (1), then, are cue-based, i.e. dependent on associative, non-inferential predictions, in agreement with our primary-trade-off: it will be worth it, in most contexts, to endorse a false alarm by reacting to a negative affordance even when there is no harmful property in the environment, rather than to omit to react to it, and incur a serious risk.

Feelings non-objectively represent an affordance through a graded appraisal. They cannot be wrong, because they do not have a propositional format and do not assert anything. However, they can be appropriate to a situation or not as a function of the success of the action that is triggered. Appropriateness of an affordance-sensing means that the impulsive action set off by it will be successful, i.e., will have the programmed outcome. In inappropriate impulsive actions, energy will be spent to no avail. It is interesting to observe that appropriateness cannot be fully understood from the viewpoint of the impulsive action system. It can only be appreciated from the viewpoint of an external viewer, or from the viewpoint of the agent herself, when able to judge what she did.

Impulsive actions are a default mode of reacting when an agent is confronted to a significant change in her environment. This change consists in a sensed affordance, which disposes an agent to react. They are the most functional choice when there are intrinsic, task-related temporal constraints favoring swift reactings (in sport, in conversation, in dangerous situations). If they are innocuous, or in case the agent divides her attention among several goals, reactings tend, then, to outnumber the other types of actions: conversational gestures, facial mimicks, expression of positive affects, and of shared negative affects. They may also be inhibited, however, when an agent can control the feeling-based representational mode by rational considerations based on beliefs and desires (in the strategic mode of action: see Section 2.3. below).

[8] Barrett and Bar (2009).

2.1.3 Inputs of Impulsive Action are Not Insulated from the Verbal System

The possibility of using language impulsively shows that it would be incorrect to take the representational system for reacting to be cognitively insulated from the verbal system of representation. Not only can adult humans entertain *simultaneously* impulsive and conceptual representations: they can also freely switch usages of each representational format. They can reactively use words that may prima facie appear to express conceptual representations. For example, an excuse can be offered by uttering a sincere, reactive "sorry", or by a strategic, conceptual, insincere, but possibly perfunctory "sorry". The two reactive systems process information and influence decisions on the basis of their own narrow range of associations and norms, while the system for strategic action, as will soon be shown, takes advantage of background beliefs and inferential reasoning to make decisions in light of a broader set of norms. *But there is no limitation concerning the range of inputs that can trigger a reaction*. For example, an agent can sense an affordance after having spent a lot of time and effort in building up a complex conceptual representation of a situation. Think of financial traders, chess players, etc. What is decisive is the mode under which they act: an affordance-sensing is non-doxastic, but it can be triggered at any point in a belief-based type of reasoning. This will become clearer below when we come to the question of cognitive reactings.

2.1.4 First Objection: Value and Time

An objection could be made against the proposed association between impulsive action and a temporal constraint of urgency: an individual might detect a mating opportunity and react to it in an impulsive, fast way. The value of his action, however, is procreation, which occurs much later in time (if it does). Does this not suggest that the time scale of an impulsive action cannot be based on fast reactivity alone? This objection, however, presupposes that feelings represent the distal goal that explains why these feelings were selected: helping agents to detect the opportunities that will serve their survival. This is not, however, in the present view, what feelings represent.[9] On the present view, the representational content of feelings is based on their informational basis, rather than on their teleological history. The appraisal which feelings make salient has a valence and an intensity that concern a present affordance. The latter is not evaluated for the good or bad consequences that might ensue later from acting on it. The short-sightedness of feelings is such that the agent only focuses on reacting to the present affordance in the restricted representational format that the feeling makes available to him.

2.1.5 Second Objection: Primitiveness

A substantive worry that some readers might have is that the impulsive actions as described above are too primitive to count as actions: they are based on associations rather than beliefs, and seem to be "wired in" by evolution. A few clarifications should hopefully dissipate this worry. First, there are strong reasons for claiming that some

[9] See Proust (2014a).

types of action do not depend on the agent's beliefs and desires. Excluding them from the realm of action because they fail to correspond to the classical concept of action as based on practical reasoning does not seem justified if they turn out to be adaptive and to systematically depend on incoming information. Second, on externalist approaches to action, agency does not require from agents the ability to articulate their reasons to act. Impulsive reactings count as actions, rather, because they are intentional, involve a form of control—can be inhibited—and have conditions of appropriateness, to which the agent is sensitive. Thirdly, the system on which impulsive actions depend does not have to be totally insulated from the doxastic system. Agents who act on an affordance-sensing may have in parallel a complex sets of beliefs and desires available for representing a given situation. The proposal is merely that, when they act impulsively, it is not their beliefs and desires that are causal, but rather their momentary affordance-sensing. In the fencer's case, for example, the agent may have a complete knowledge of the rules and of the winning strategies in particular circumstances. But he cannot make a winning point if he lets his own action be guided by these considerations. Impulsive action, then, is not a failed attempt to act on the basis of one's own practical reasoning. It has its own role to play in agency, and is cultivated as a way of making the best practical decisions even when time is too short for the agent to form explicit, concept-based judgments.

2.1.6 Comparison with Alternative Theories

This analysis has the merit of accounting both for the immediate reactivity and for the flexibility of an impulsive action. Jesse Prinz (2004) offers a theory which accounts for flexibility, but not for immediate reactivity. He mentions in passing Gibson's affordances (Gibson 1977) as providing a relation between percepts and behavioral responses. He does not take them, however, to "impel actions directly". They, rather "instigate the search for appropriate actions". In other words, emotions are not "actions commands".[10] Impulsive actions are not envisaged as responding to temporal constraints. On our view, in contrast, impulsive acting is part of what sensing an affordance at a place is. As an embodied, dynamic representation, with both a registering and an imperative component, the representational structure is designed to deal simultaneously with perceptual input and with behavioral output. This dual role makes our affordance-sensings a case of what Ruth Millikan called "pushmi-pullyu representations" ('PPRs'),[11] which, being "more primitive than either purely directive or purely descriptive representations", combine both directions of fit. Millikan, however, did not try to characterize further the semantic content of PPRs, nor did she relate their dual role to their having a feeling-based content.

Andrea Scarantino's motivational theory of emotions similarly accounts better for the flexibility of impulsive actions than for their urgency. As noted above, he accounts for the latter by emotion-based control precedence and rejects a belief-desire account, as our own proposal does. He offers, however, a teleological account of what emotions represent as a function of their "abstract" core relational theme. His notion of PPR,

[10] Prinz (2004), p. 228. While conceding, p. 229, that valence markers are "internal commands to sustain or eliminate a somatic state by selecting an appropriate action", Prinz considers emotions to be perceptual, rather than agentive.

[11] Millikan (1995). This feature of impulsive action is also discussed by Bar-On (2013) and by Scarantino (2014).

accordingly, refers to the dual properties that he attributes to emotions, namely that they consist in dispositions to act, and that they represent facts by having the function of being elicited by them. As noted above, however, this teleological detour affects the informational directedness of the specific affordance-sensing and its associated reaction. PPRs loose their dual character, if a disposition to act is not represented *within* an individual affordance-sensing.

2.2 Routine Action

Routine actions are another default mode of acting, instantiated when an agent has to deal with habitual, recurrent situations, such as driving, preparing food, adjusting posture, using tools, playing videogames, practising musical instruments, etc. Although they have some structural similarity with impulsive actions, routine actions have a different relation to time: their own way of being triggered "here and now" is based on recurrent schemas rather than on sudden changes in the environment. They also fail to have the same urgency.

2.2.1 Similarity with Impulsive Actions

Let us first emphasize, however, their similarity with impulsive actions. Just as the latter are performed "in the grip of an emotion", the former are performed "in the grip of a habit". As is the case for impulsive actions, practical reasoning (which is time- and resource-consuming) is not called for. This absence is revealed in various functional characteristics of routine actions also present in impulsive actions: 1) they respond shortsightedly to one feature of the situation that exerts a form of control precedence on the agent's disposition to act. Agents are thus commonly led to deviate from their prior intention to act, (e.g., ceasing to add sugar to their morning coffee) by unwillingly following an habitual response sequence. 2) While in the grip of a habit, agents tend to overlook unwanted consequences of their behavior in the present context. 3) Routine actions can be performed automatically, with little focal attention, and sometimes with no conscious awareness (as the case of unconscious driving behavior strikingly demonstrates). They can also be performed, however, as part of strategic, volitional actions. 4) Their appropriateness or inappropriateness relates to their having or not having their expected, usual outcome.

Taken together, these characteristics suggest that routine actions are also meant to respond to present opportunities, and to be based on a restricted part of a perceived situation: these opportunities for acting might again be represented by a non-propositional "pushmi-pullyu" affordance-sensing. Emotion-based affordances, however, differ from the utility-based affordances involved in habitual behavior. Phenomenologically speaking, a habit consists in perceiving familiar objects as allowing us to reach a given end, in an embodied, nonconceptual way. The dispositions to act associated with these affordances, then, consist in instrumental, pragmatic schemas, through which the agent feels immediately and unreflectively "drawn to act in a certain way", [12] as a function of the intensity and valence of the sensed utility-based affordance.

[12] In the terms of Dreyfus and Kelly (2007).

Such a disposition can be strategically controlled when conditions of intensity and valence in the affordance-sensing allow it. For example, a tired agent on a formal visit should resist sitting down uninvited in an armchair even though strongly tempted to do so. If very tired, he may yield to the temptation. The possibility of controlling one's own routine actions raises interesting problems concerning how the control switch is operated. We will address this question in the last section.

2.2.2 Differences with Impulsive Actions

A first difference between routine and impulsive actions has to do with the type of feelings involved. On our view,[13] feelings are not limited to affective reactings. Arguably, the feelings that constitute utility-based affordance-sensings in the case of routine actions are not of an affective nature. Their evaluative role is to represent opportunities relative to one's own (and to others') well-being and utility. Relevance to well-being extends to bodily conditions (thirst, pain, exhaustion), availability of preferred entities of all kinds (food, partner, wealth, and social status). Relevance to utility covers the multiple cases in which we use tools, furniture, electronic devices, etc. and select them as the best we can get, i.e. those that score highest on the affordance concerned.

A second difference has to do with the contrast between recurrent, expected and non urgent, versus contingent unexpected, and urgent opportunities. The recurrent opportunities inspiring routine actions may thus be of a sensory, proprioceptive (feeling thirsty, cold, etc.), social-affective, (feeling isolated, honoured, etc.) or instrumental nature (objects can be sensed as more or less useful for performing one's recurrent activities). As a consequence, in contrast with impulsive actions, the range of routine actions *fluctuates as a function* of time of day, context, and cultural environment. Routine actions being a response to utility and well-being affordances, they tend to evolve over time as a function of the current incentives for possessing and using various types of triggering objects. This easel may no longer elicit in its possessor an intention to paint, this piano to play, this camera to take pictures. New objects are now involved in the affordance-sensings of a given agent.

A third difference, then, is that, given the entrenchment of objects and practices in our socioculturally determined way of life, routine actions are more culture-specific than impulsive actions. Another related difference is that they spread from one agent to another, by a process akin to the epidemiology of representations, i.e., by the mere strength of the example of observing another's affordance sensing. Agents have taken the habit of automatically opening Facebook when they get back home, or wearing their earphones while waiting on line, by learning from others the corresponding affordances.

2.2.3 Transitions from Impulsive to Routine Actions

Note, then, the interesting process exemplified in the latter kind of case, through which an affective feeling is replaced by a mere instrumental affordance-sensing, with the associated transition from an impulsive to a routine action. Affective feelings have first prompted agents to impulsively acquire a device, have a romantic relationship, etc. Over time, the affect-based affordance becomes an affordance of the recurrent, utility

[13] See Proust (2014a).

kind. We will see that this kind of transformation also frequently occurs from the strategic to the routine actions, but never conversely.

2.2.4 Objection: Valence, Intensity, Reward

It might be objected that feelings of well-being or of utility cannot consist of evaluations based on intensity and valence gradients. Picking-up my toothbrush does not respond to an intensely positive evaluation of what it can afford. One might claim, rather,[14] that routine actions merely respond to a learned association between a given perceived stimulus and the executive representations for dealing with it. An executive representation, also called "motor representation", is defined as a non-conceptual representation that can be activated in multiple ways, in order to shape, guide and monitor a given action until completion. Granted, the toothbrush and the armchair are opportunities for acting routinely. But this only involves associative conditioning, where perceived pre-potent stimuli are able to activate a given executive representation.

The main arguments offered in favor of this account are that routine actions are egocentric, shortsighted, and automatically activated through typical motor patterns. There are additional functional reasons, however, for resisting these arguments. First, the features of egocentricity, shortsightedness and automaticity also characterize our analysis of routine actions above, in terms of utility-based affordance sensings. Motor representations inherit these characteristics from reactive (i.e., affordance-based) actions, and do not constitute actions in themselves. Executive representations are involved in selecting this grip rather than that that one to hold a pan or a racket. They also have the right format for expressing the dynamic interactions between body and goal. They do not have the function, however, as affordance-sensings do, of detecting and responding to recurring opportunities and resources by triggering an action. Let us suppose that two objects are available for performing a routine action. Shall I pour a glass of red or white wine? Shall I use a nail or a screw to fix a frame? Motor representations have no say in this: the distal targets of action have already been selected when they come into play.

In Reinforcement learning theory, habit-driven, also called "model-free", learning, is shown to occur against the background of a pre-existing motivational state, such as hunger or satiation. In conformity with our feeling-based representation, learning consists in detecting and assessing a utility on a graded scale for its valence and intensity. A recent important finding[15] suggests that routine actions differ from controlled actions by being immediately insensitive to the extended utility of the action (i.e., to its specific consequential outcomes). They are immediately sensitive to "stimulus–action values", rather than to the further, long-term consequences of this action. For example, hungry rats recently trained to press a lever to obtain sucrose pellets will reduce their lever-pressing when the outcome has been devalued (pellets now have an aversive taste). After a prolonged training, however, when lever-pressing has become habitual, rats are insensitive to the modified utility, and persist in lever-pressing as long as pellets are provided. [16] Similarly, addicts find it difficult to change their addiction

[14] See Pacherie (2002, 2011) & Proust (2003).
[15] Lee et al. (2014).
[16] See Niv et al. (2006), p. 377.

habit even when the outcome has ceased to be pleasurable, and in spite of dramatic consequences for their social life. This form of "short-sighted" motivation is also part of the concept of affordance, which only depends on valence and intensity of the disposition to act, and has no access to further motivational considerations.

To summarize: Our discussion so far has allowed us to distinguish two kinds of affordance-sensings, corresponding respectively to affective and non-affective types of feelings. Affective feelings trigger impulsive actions. Non-affective feelings generated by habits also trigger a routine action by non-reflectively motivating the agent to act in a particular way. Routine actions play the most prominent part in human behavior, plausibly because they represent an efficient compromise in our temporal trade-off in stable environments. They are rapidly performed, and at a lesser cost, than reflective actions.

This analysis helps us address our initial question concerning the representational cause of action. In the two cases discussed, a feeling-based evaluation,—an "affordance-sensing"—, has the role of motivating and triggering an action. It differs from the classical concept of a "minimal intention", or "intention in action", by depending on affect or reward-based evaluations of what a given situation affords to the agent, rather than on a conceptual representation of the goal or of the specific ways for attaining it. Being a pushmi-pullyu representation, an affordance sensing both represents the agent's relation to an opportunity and the action that the latter motivates.

2.3 Strategic Action

Given its relevance for rational action, i.e. action based on practical reasoning, what we are calling "strategic action" (SA) has been the main focus of philosophers. It has been articulated in terms of prior intentions formed through deliberation.[17] The ability to coordinate one's thoughts and actions in relation to an internal goal, formed independently of present solicitations, is a crucial adaptation. SA could also be named "rationally controlled action", because it is based on large sets of propositional beliefs and desires that can be recombined, checked for their practical value and coherence, shared with other agents, and because it enables agents with planning and explicit rationalization capacities. Contrary to a general assumption, strategic actions, and inhibition of habitual responses, have been found in non-humans.[18]

Our present task, again, is to identify the function of such a type of action in terms of our trade-off between needs, resources and timing in decision-making. How might this trade-off have reversely engineered strategic action? The first striking feature is that, in contrast with impulsive and habitual actions, no control precedence, no prioritized disposition to act are involved. Accordingly, strategic actions are not influenced by present cues, nor are they typically or exclusively concerned with immediate goals. Furthermore, they do not develop automatically and they are not controlled by habitual reward feedback. This set of negative features clearly depends on specific ecological constraints. Experimental evidence shows that stable contexts invite habitual actions. Strategic actions, in contrast, are elicited in unstable, changing contexts (Ouellette and Wood 1998). On the positive side, performing them requires conscious attention, and

[17] Searle (1983), Bratman (1987), Mele (1992).
[18] Lee et al. (2014), p.695.

most importantly, a more or less temporally extended computational, reflective episode devoted to planning. Once formed, the decision to act at a later time—following planning—is temporally stable. In other terms, the representational contents of prior intentions are no longer deliberated about.

2.3.1 Contextual Instability

What kind of instability in contexts of action might have contributed to selection of individuals with a capacity for strategic action? A higher variance in the availability of food or territory might be among the contributing pressures. In the human case, the growing complexity of social interactions in larger groups, including the rise of cooperation of individuals within and across social groups, might have generated variability in procedures and uncertainty about their outcomes. As we saw above, habitual action is not equipped to track the modifications of utility, nor does it involve representation of the long-term consequences of a decision. Appreciating flexibly the dynamics of utility of one's actions within and across contexts, and predicting the long-term consequences of a plan, however, are decisive elements for efficient planning. For example, evidence predicting food shortage, climate change, major epidemic risks, rationally recommends taking strategic actions meant to overcome the problems even before they appear. Having only habitual actions in one's repertoire, then, becomes a major obstacle to individual and collective survival. An action system allowing agents, long before acting, to acquire information relevant to future well-being, to come up with possible alternative lines of action, identify new *and* accessible goals, is likely to better confront future risks and to generate larger benefits than an exclusively habitual action system could do.

Let us suppose, then, that our third mode of acting is meant to cope with contextual instability. How then, should action be organized to respond to this constraint? Let us discard all the adaptations already present in habitual and impulsive actions: these actions are able to adjust to the specific characteristics of the environment. They can select the proper motor representations for their performance to be prompt and efficient. They are able to terminate action when the goal is attained. Habitual actions are also sensitive to their most salient present utilities when selecting a given end.

Contextual instability, however, raises several time-related puzzles relevant to reverse engineering. First, reasoning about uncertain events. Second, imagining scenarios for acting, and selecting the most reliable given present uncertainty. Third, securing the coherence of one's plan until it is executed. These three requirements have exerted selective pressures in favor of a capacity to perform the corresponding mental actions, which as will be seen in the next section are always embedded in strategic actions. Their function, among others, is to assess uncertainty concerning perceptual, memorial, or motivational properties associated with success in a given distal task. It is also to imagine alternative solutions in problem solving, and to appreciate the coherence of sets of beliefs.

Our present focus, for now, is on the properly executive part of the type of action we characterize as strategic. The challenge for this type of action is to secure execution *at a point in time later* than the time the plan was initially selected, while allowing contextual adjustments to be made if necessary. This is a serious achievement, from the viewpoint of the evolution of action, as until then, execution was triggered by recurring events or situations, rather than by an endogenously maintained

representation. The crucial adaptation consists in evolving a form of prospective executive memory, extending further than a single context, across the entire episode that the intention to act defines. It is a form of prospective remembering, where the recollection of the action that has been planned can rely on a sensory trigger such as "when event *e* is perceived, then do SA". It may also rely on endogenous attention periodically bringing the intended action to mind.[19] Functional neuroimaging suggests further that there are two distinct components to prospective memory: the maintenance over time of an intention, on the one hand, is a lateral prefrontal matter, while the execution of the planned intention activates the thalamus. [20]

2.3.2 Temporal Trade-off in Physical Strategic Actions

Interestingly, keeping an intention active in prospective memory slows down the cognitive activity that is performed during this time.[21] This suggests that the proportion of strategic actions in overall agency should be kept at the lower level possible compatible with the importance of the stakes (risks and benefits) affected by context instability. A related prediction is that strategic actions should tend, whenever possible, to include routine rather than strategic actions as their subcomponents. Taking advantage on automatic reactions should lessen the resources needed to engage in action control, while also elevating the speed of execution. These two observations can be assessed on the basis of experimental evidence, concerning respectively the selection of a type of action, and the hierarchical organization of strategic actions.

Selecting a Mode of Action If our first observation above is correct, some arbitration must occur between a disposition to perform habitual, cheaper routines, and a more demanding planning-based control. How can the predicted reliability of each system be compared, in order to pick the most reliable given its cost for obtaining an expected gain? Recent experimental evidence[22] suggests that such arbitration takes place subpersonally, based on comparing the mean Reward Prediction Error in the habitual system, with the mean State Prediction Error assessed in the "model-based" strategic action system. The neural implementation of the components of the arbitration process has now been identified.[23] In agreement with our hypothesis, the transition between systems incorporates a bias favoring habits over strategy in action control. This bias is reflected in the phenomenology of action, where feelings of effortfulness and ease of performing have an opposed valence, and dispose agents, all things being equal, to prefer habit to strategy. We will see that mental actions present the same bias.

Selecting the Subcomponents of an Action Neuroscience allows us to address the second issue, by parsing out the various types of cognitive control involved in a task comporting various sub-tasks of perceptual discrimination. [24] Analysis of the neural

[19] McDaniel and Einstein (2000).
[20] Burgess et al. (2001).
[21] Smith (2003).
[22] Lee et al. (2014).
[23] Notably, evidence for a comparison signal reflecting the difference in reliability between the two systems has been found in the rostral cingulate cortex. Lee et al. (2014), p. 693.
[24] Koechlin et al. (2003). See also Fuster (1989), Grafman (2002).

activity reveals a cascade structure, where a higher structure holds over time a representation of the distal goal and of its main steps. What is noteworthy for our purpose is that the various control structures activated by this task are organized hierarchically, *as a function of the temporal scope of each control activity*. The activations revealed by fMRI indicate that medial and lateral orbito-frontal structures monitor coherence of action with one's own represented self and long-standing values.[25] The representation of one's long-term goals is maintained over time in the rostral part of the lateral prefrontal cortex (LPFC), an area whose role in prospective memory was emphasized above. This structure controls in turn the activity in the caudal LPFC, which monitors the various sub-episodes—temporally shorter context-dependent activities—relevant to attaining the goal. Finally this contextual structure in turn controls the motor responses that are triggered and monitored "here and now" in the premotor cortex. This cascade model of action control, then, reflects the temporal structure of the developing action, through the representations that underlie executive control.

The second expected consequence of our primary trade-off, then, seems to be compatible with the temporal dimension of the cascade structure just discussed. Strategic actions tend to divide the tasks into simplified, shorter sections, whose execution is performed by slave systems. Characteristically, there is increased short-sightedness when descending levels, and hence, a decrease in cognitive resources. The lower structures are, then, faster in sending their instructions and monitoring them than the structures above them.

This in turn entails that, when cognitive resources need to be spared, for example when several tasks need to be simultaneously attended, cognitive control of actions will tend to be limited to the lower levels in the cascade, namely the contextual and the sensori-motor levels. In other terms, agents will turn to habitual modes of action.

2.3.3 Categories of Intentions

Taking our outside-in perspective on action, then, suggests an alternative way of addressing the Fundamental Problem of the theory of action. The concept of intention as classically understood can be cashed out as three types of causal-representational event: emotional affordance-sensing, utility affordance-sensing, and planning. The first two are represented in non-propositional, non-conceptual, gradient-based terms involving valence and intensity. The last type, which corresponds to the classical prior intention, crucially involves beliefs, desires, and the mental action of planning to act, along with mechanisms of prospective memory.

Some readers may be surprised by the fact that, on the present view, motor representations do not qualify as motor intentions. Should not such intentions be postulated, however, in order to account for the selection and guidance, at a finer time scale, of the specific type of movement to be used in a given context?[26] In response to this observation, however, motor representations could qualify as intentions if they had their own independent level of control. As we have seen, however, the visuomotor level, being at the bottom of the hierarchy, is a slave mechanism for higher levels of

[25] See Damasio et al. (1994).

[26] For a defense of motor intentions, see Pacherie (2006, 2011).

control. Contextual control, when conducted independently from an overarching strategic intention, belongs to habitual control, and then depends on affordance sensings. Such habitual forms of intention rely on motor representations for enacting them. As Koechlin observes, "the premotor cortex integrates all these [level-specific] signals with those of the stimulus to decide how to act".[27] A decision to act is driven by the intention located at a higher level; either at the highest level, sending commands to all the lower-level subsystems involved in the cascade, or at the immediately higher level, where habits deal with contexts, in repetitive but motorically flexible ways. On this view, intentions are formed at the highest level in the cascade that they happen to activate. The lower level of the executive hierarchy consists of motor adjustments, not of motor intentions.

The evidence for a hierarchy of control subsystems, combined with our principle of economy, suggests that an agent should form an intention on the basis of two types of constraint: i) that this intention is in the agent's repertoire; 2) that the executive properties of this intention-level are adaptive for the situation in which they have been selected. Affordance-based intentions are needed for many situations where prompt action is needed. Ordinary chores rely on habitual reactings, which, as we saw, can be "upgraded" into context-based strategic actions. Other situations will require careful pondering of the pros and cons, and possibly engage representation of the episodic constraints attached to the present performance (for example, visiting a place with different food norms from one's own), of one's own moral norms and one's own view about oneself as a person. In this kind of exceptional case, higher-order intentions will be formed for acting strategically.

Let us conclude our remarks so far, by summarizing when, from the viewpoint of our primary trade-off, a given type of action, with its associated intention, should prevail over the others. What makes an impulsive action appropriate is that it is performed in an automatic, fast and effortless way, in order to deal with fitness-relevant, unpredicted events. What makes a routine action appropriate is that it is performed quickly and effortlessly in order to take advantage of recurring opportunities in stable contexts. What makes a strategic action appropriate is that the agent's beliefs, norms, and preferences are used to confront situations involving changing contexts with variable, uncertain utilities.

2.3.4 Explaining Action Failure

The cascade model presented above suggests that action failure may have two main general sources. One is that the agents have misrepresented the type of context in which their action is taking place. They have failed to realize, for example, that the contexts of their action are changing over time, or that there is an immediate threat they should react to: they will opt for a routine action when an impulsive reaction would have been adaptive. Or they have chosen to follow habit in a situation that required strategy.

A second main type of failure is caused by disconnections of the levels in the cascade. A prior episodic intention fails to appropriately guide contextual decisions. For example, the agent wanted to stop off at a supermarket on his way back from office, and simply takes his habitual route back home, which does not go that way. More

[27] Koechlin et al. (2003).

severe disconnections occur in patients with schizophrenia. They may fail to recognize their strategic intentions as their own, for lack of an appropriate relation between the episodic, contextual and sensory-motor levels that are articulated in their strategic actions.[28]

We now need to explore whether our primary trade-off also applies in the case of mental actions.

3 How our Primary Trade-off Shapes Mental Actions

By a "mental" (or "cognitive") action, is meant an attempt to control one's own perception, memory, reasoning, or other cognitive function, in order to attain a given informational outcome. Although the term of "cognitive control" is sometimes used to refer to any form of control operated within a cognitive system, and in particular, to the control of one's ordinary actions, we will here restrict its application to the control of one's cognitive actions, i.e., to such cases as trying to *discriminate* a shape among distractors, to *remember* the name of a place, to *solve* a problem, to *learn* a list of names, or, crucially, to *deliberate* about what to do. As indicated above, mental actions are usually embedded in physical actions. For example, assessing one's ability to remember the items in one's shopping list is a precondition for pursuing the shopping even though the written list is not presently available. Planning is a complex mental action which, as seen above, is a precondition for acting strategically. As will be shown shortly, however, many other cognitive actions are also involved in physical actions, whether strategic or habitual. An interesting question is whether a mental action has to be habitual (or strategic) to be embedded in an habitual (or strategic) action. To address it, let us first come back to our methodological guideline. Has the primary trade-off also shaped types of mental actions and hence, types of proximal events triggering them?

3.1 Impulsive Mental Actions

By parity with impulsive physical actions, one might speculate that impulsive cognitive actions are also a specific kind of disposition to act having control precedence, as part of a specific affordance-sensing. This speculation turns out to be correct. To provide evidence for it would require more space than is available here. Hence, we will summarize arguments that have been presented in detail elsewhere.[29]

A typical case of an impulsive mental action is the rapid decision associated with having a feeling of knowing (FOK). One senses that one can remember a presently elusive name, and immediately decides to try to recall it here and now. Or, conversely, one senses no affordance for remembering the name, and immediately stops trying to remember it. In both cases, these feelings are subcategories of "feelings of fluency". These feelings evaluate the ease of processing associated with a cognitive task, either one to be performed, or one already performed. They are, then, a conscious expression of the fact that the primary trade-off is satisfied for a given combination of urgency, low cognitive load and moderately low expected reward.

[28] See Proust (2013), Chapter 12.
[29] Proust (2013).

Noetic feelings play a fundamental role in our lives because they produce fast but reliable evaluations about likely outcomes. Some lead us to impulsively perform or reject the cognitive task at hand. Others retrospectively recommend acceptance of an outcome. The feeling of being right, for example, is experienced when agents have retrieved what seems to them to be the name they wanted. Having a noetic feeling bears close similarity to having a feeling of dangerous or of pleasurable opportunity discussed in Section 2.1. It also denotes a reactive, subjective experience with a distinctive embodied phenomenal quality associated with the pleasant relaxation (or unpleasant contraction) of dedicated facial muscles. When such a feeling is experienced, the agent senses a "knowledge affordance" with a certain gradient of intensity and in valence. The latter is often verbally described as the subjective sense of confidence in one's ability to retrieve the name. The affordance comes with a disposition to immediately perform (or reject) a cognitive task or to immediately accept (or refuse) to act on its outcome.

Cognitive affordances, however, in agreement with the primary trade-off, motivate a fast and modular, i.e. inexpensive response. Being sensed in a non-doxastic way, they do not require a capacity to conceptualize, and even less one to report them. Just as agents can feel afraid without having the concept of fear, they can have a FOK without having the concept of knowledge, or of uncertainty. This is our claim. What arguments can we bring up to support it?

3.1.1 Animal Evidence for Cognitive Affordance-Sensings

First, evidence for this claim comes from comparative psychology. Rhesus monkeys, working in experimental labs of comparative psychology, are able to assess their perceptual and memory affordances (even when they do not receive any trial-by-trial reward).[30] What kind of "cheap" feedback, then, do monkeys and humans use when they experience a reactive feeling?

3.1.2 What is the Information Being Used when Experiencing Noetic Feelings?

Our second argument is the response that experimental psychologists have been able to offer to this question. Noetic feelings do not depend on analysing the semantic contents of memorial or perceptual states. They are, rather, "mediated by the implicit application of nonanalytic heuristics, relying on a variety of cues."[31] These cues are "activity-dependent", which means that they are extracted from the pattern of cognitive activity occasioned by the present task. Activity-dependence is associated with an important type of contextual flexibility: activity-dependent cues tell agents where they stand, here and now, relative to their cognitive dispositions in a given task. Neuroscientific research confirms Koriat's claim. The relevant cues have to do with the dynamics of information processing: with its onset, with the comparative amount of neural activity favoring alternative responses, and with the time needed to converge on a threshold value..[32]

[30] See Couchman et al (2012).
[31] Koriat (2000).
[32] For example, the neural activity recorded in rats' orbitofrontal cortex when attempting to categorize olfactory stimuli was found to correlate with their predictive behavior (i.e., accepting or rejecting a task trial). See Kepecs et al. (2008). Similar patterns have been found in other species, including humans. For a discussion and review of the literature, see Fleming et al. (2012), and Proust (2013), pp. 99 sqq.

As in the case of physical reactings, then, cognitive affordance-sensings are evaluative rather than assertive; through their felt intensity and valence, they guide epistemic decisions in a quick and economical way. In contrast with physical reactings, however, the disposition to act associated with a feeling is restricted, in cognitive control, to the acceptance or rejection of a current task (when the feeling is produced before the task is executed), or to acceptance or rejection of the output of a task (when the feeling is experienced after the task has been executed). Even in this restricted class of cases, the agent's cognitive decision is made as a function of the valence and intensity of activity-dependent feelings and cues. This justifies the claim that cognitive reactings are types of impulsive mental actions, just as physical reactings are types of impulsive physical actions. Their function is to respond to the economical constraints that our primary trade-off imposes on action: Act cheaply and fast, but as reliably as the stakes recommend, in a given new context.

3.1.3 Fitness Value of Cognitive Affordance-Sensings

Our third argument emphasizes the added fitness value of having impulsive cognitive actions in one's repertoire. In non-humans, feeling-based sensitivity to cognitive uncertainty has been hypothesized to enable an enhanced efficiency in competitive foraging.[33] Highly competitive, semi-nomadic macaque monkeys seem to be eager to predict what they will find. In contrast, arboreal capuchin monkeys live locally rather than nomadically in environments providing ample resources of various kinds of food. The former, but not the latter, prove able to evaluate their own ability to perceive or remember.

Even before they are able to solve a false belief task, young children can rely on their feelings of fluency to guide such impulsive actions as forming preferences about their informants, attending to objects of interest, determining when they can reliably perceive or remember an item etc.[34] Humans, more than any other species, are continuously screening their cognitive affordances in order to retrieve items from memory, optimize perceptual input, control learning, redirect cognitive effort, select problems they can solve, etc. There is no area of intellectual endeavour that does not need to rely on feeling-based, impulsive epistemic decisions.[35] Because they are fast and frugal, they are used as additional means of managing the informational states that are engaged in the other two categories of mental actions as well.

3.2 Routine Mental Actions

In contrast with the other primates, controlling and monitoring one's informational states (i.e. performing mental actions) is no longer, in humans, merely a matter of feeling-based affordance-sensings. Under the joint influence of language learning, school training, and, more generally, social learning, every one of us has a huge repertoire of cognitive routine actions.

[33] Carruthers and Ritchie (2012).
[34] See Balcomb and Gerken (2008), Bernard et al. (submitted).
[35] See the review by Greifeneder and Unkelbach (2013).

3.2.1 Temporal Trade-off in Routine Mental Actions

These routine mental actions satisfy our primary trade-off, in the specific case of stable contexts involving recurrent opportunities. Habitual actions, being repeated over time, have the cognitive property of being fluently processed and carried out. With little effort, then, cognitive routine actions can produce quickly and effortlessly a given cognitive outcome whose nature will depend on the opportunity. Familiar examples are looking up a number in a phone book, a word in a dictionary, reading a newspaper at breakfast, verifying the time of one's next appointment, and most of all, asking questions, i.e., trying to extract information from others.

Again, according to our theory of feelings, responding to such opportunities does not have to depend on concept-based inferences. As we have seen, physical affordances can be sensed independently of their being conceptualized as such. One can speculate that, similarly, cognitive affordances may be involved in an agent's habitual behavior without using the full-blown theory of knowledge acquisition that the agent may have independently acquired.

The examples above are bona fide cognitive actions: an informational goal is felt to be within reach, or, in our words, a cognitive affordance is sensed, which, in the relevant instrumental context, immediately triggers the associated behavior. Yet the information is gained unreflectively, through a habitual, automatic procedure, rather than through an intention to act based on practical reasoning.

3.2.2 Instrument-Aided Cognitive Habits

As in the parallel cases of habitual physical actions, agents can sense a cognitive affordance associated with a familiar type of object (agendas, phone books, dictionaries, calendars, computers), an informant, or a place (a museum, a zoo, a library). Given their present context, and their history of cultural or cognitive practices, agents may feel "immediately and unreflectively drawn" to use one of these objects, as a function, again, of the intensity and the valence associated with this affordance sensing in that context. For example an outdated phonebook will have lost much of its utility, as will a regularly unreliable informant. As in the case of physical actions, then, agents engaged in a cognitive routine are sensitive to its present convenience, not to its long-term potentially harmful consequences. Habitual use of externally mediated cognitive routines, for example, might considerably reduce agents' individual cognitive abilities without their being aware of this.[36]

3.2.3 Mental Tools and Cognitive Habits

All the examples of routine cognitive actions given above involve objects or persons. A number of cognitive affordance-sensings, however, also related to utility, involve individually engineered mental tools for retrieving, memorizing, or learning. These routine actions are often based on internalized spatial or auditory features of the information to be encoded or recovered. Think for example, of the alphabetical search one uses to retrieve a name, or the Greek method of loci for efficient encoding in

[36] See Proust (2014b).

memory.[37] A third example is repetitively rehearsing some material to memorize it. There are many more procedures that agents invent as a function of their own recurring needs.

A particular additional type of case deserves special scrutiny. Individual subjects or institutions may often deliberately (i.e., strategically) work at turning either impulsive or strategic actions into habitual ways of acting (with their respective affordance-sensings). Feelings mimicking those produced impulsively can, through a long controlled process of self-transformation, be produced within habitual—not strategic—actions. We will postpone discussion of such cases to the Section 3.3.3.

3.3 Strategic Mental Actions

In Section 2.3, contextual instability was claimed to have been the factor that led to a need for new constraints on agency. Physical strategic action emerged as a solution enabling a temporal trade-off in a new type of equilibrium: more cognitive resources must be collected over time and applied to prepare the action. Thus more time will be needed to form an intention to act. Execution is also typically slower: Aside from the fact that it is designed to take place at any point in the future, the action will be executed more slowly, because of its lack of fluency, and the need to pay attention to consequences throughout performance of the action. Expected benefit and risk avoidance, however, have a much more extended application, in a world that has much more to offer—and to lose—than the stable, repetitive world of habit. In spite of the cost, SA occasionally furnishes an efficient compromise relative to our primary trade-off.

As also mentioned in Section 2.3, reliance on practical reasoning, anticipation of consequences and control of prospective memory make mental actions a necessary ingredient in physical SA. As will soon be clear, it is in these roles that mental actions themselves must become strategic. Impulsive and habitual mental actions, however, also contribute to SA, in particular in assessing the memorial uncertainty associated with preconditions of action, and with the perceptual uncertainty pertaining to cues relevant to action execution.

In order to execute successful strategic mental actions, agents must also have structured and combinable propositional knowledge concerning the environmental context and the associated constraints for acting strategically. They must in particular be responsive to epistemic norms beyond fluency, which is the only norm to which noetic feelings are sensitive. Additional epistemic norms, such as accuracy, plausibility, and exhaustiveness, are conveyed through language and learned by social observation and formal education, rather than made immediately salient through dedicated feelings. Examples of the variety of epistemic norms available will be given below.

3.3.1 Planning for Physical Strategic Actions

Planning is a strategic *mental* action that governs a strategic *physical* action. It allows agents to determine the specific goal to be attained among alternatives, to deliberate about the most efficient ways to attain it, and to consider the consequences of each

[37] This consists in associating the items to be remembered with specific familiar physical locations, for example the rooms of the agent's house.

instrumental step, as well as those of the world change that the action aims to introduce.[38] Preparation of the action must occur at a temporal distance from the action. The optimal temporal distance is determined by considering the context where the action either must take place (when it depends in part on other agents or external factors), or is most likely to succeed.

Planning qualifies as a strategic mental action because its goal is to epistemically and conatively control and monitor the successive steps that are conducive to the goal. Many different epistemic appraisals may be required in the course of a planning episode. First, one must estimate how *correctly* and *exhaustively* the plan takes into account the beliefs *relevant* to the success of the action in a given context. Secondly, although the details of a plan are typically left relatively unspecified, it should be assessed for the *coherence* between the means envisaged and the ends pursued. Thirdly, it should be monitored for its *consistency* with the other plans of the agent supposed to be carried out over the same period (and with the other agents supposed to cooperate with him), as well as with contextual properties. Finally, it needs to be evaluated with respect to its assessment of the *uncertainty* of some of its most salient conditions of success. On the conative side, planning also monitors the motivational resources available to move forward to the goal.

Although planning indirectly has an instrumental value, in that it is embedded in a physical SA action that targets a reward, planning thus also has its own intrinsic, epistemic value. Our basic trade-off suggests that a planner will subcontract, whenever possible, part of this monitoring to the impulsive action system. For example, agents will trust their feelings of knowing or their feelings of consistency rather than engage in costly analytic procedures to select various sub-plans. Planning will also rely on routine actions for a major part of the execution of strategic physical actions—those details that are left unspecified are those that agents will feel disposed to do out of their utility-based affordance-sensings (remember that habits themselves are fluent action programs). Feelings of opportunity, then, are smoothly hierarchically integrated into a conceptual representation of the distal goal when the time comes to execute the plan (as discussed in 3.2.2).

If planning is a mental action that is a constitutive ingredient in physical SAs, can it also be constitutive of mental SAs, including planning to plan?

3.3.2 Planning for Mental Strategic Actions (MSA)

SA for Acquiring Cognitive Competences Some mental actions have as their goal the acquisition of long-term epistemic competences through a specific training or educational process. They are strategic, first, insofar as the final state they aim to bring about will occur later, sometimes years after the plan was first conceived. Second, they involve deliberating about the best ways of acquiring this competence, in terms of financing, schooling, priority given to the goal, etc. Such a deliberation involves both conative factors (motivation, preferences), and cognitive elements (beliefs about the means available, etc.). Planning one's future, then, is not only an epistemic matter. In the hierarchical conception of strategic action sketched above, one's own represented self-identity and one's own long-standing values will contribute to determining one's professional choice. Our time-resource-benefit trade-off, however, also plays a major

[38] For a concept of planning compatible with a bounded rationality, see Bratman (1987).

role in adjudicating one's own priorities, given the present and predicted properties of the context of our strategic planning. Planning to become expert in a given field, in art, technology or science, etc., is only worth it if agents are likely to attain their respective cognitive goals, and, furthermore, if these cognitive goals help them gain other forms of reward that they could not have obtained otherwise. Although planning to act breaks away from routine action, it will economize resources by relying on the latter whenever this is compatible with its ends. Note that, in certain social groups, becoming an expert at X is not planned, but rather executed because of a family habit, with no deliberation involved. In such cases prioritized dispositions to act extend over longer periods; they require, however, exceptionally stable contexts to be implemented.

Mental SAs for Attaining Collective Epistemic Goals Strategic mental actions also play a prominent role in collective planning to attain given epistemic goals. For example, evidence must be collected and carefully evaluated to fight a pandemic. Although this mental action is supposed to serve a distal physical goal, it also involves an epistemic strategy, that requires planning and deliberation. For example, when collecting evidence, the following may be important: scrutinizing the *reliability* of a given source, the *coherence* between various reports, appreciating the *probability* of error in observing the phenomenon, the *relevance* of given local usages to the spread of the disease, the *exhaustiveness* or *incompleteness* of the data collected, etc. Discerning which mental actions are required for attaining the distal goal involves a reasoned, analytic selection of the various sources of uncertainty that need to be monitored and reduced. Here again, this type of SA needs to find a compromise between the urgency of the problem to be solved and the epistemic resources to engage in order to find an efficient solution.

There are obviously more cognitive SAs than those listed above. These two forms of epistemic planning, however, represent two poles in mental agency, concerning respectively the cognitive shaping of agents and the selection of the most relevant epistemic procedures available to reach a given epistemic end.

3.3.3 Planning for Acquiring Habitual Dispositions

Although it may seem paradoxical, some cultures have invited agents to shape themselves by *strategically* attempting to replace their existing habitual dispositions to act with new ones, which will have higher value in terms of personal excellence and morality than the old ones, but will still be exercised in the automatic and spontaneous way characteristic of impulsive and routine actions. This case is interesting for us because it takes as its strategic aim that of reducing the recurring need to form costly, time-consuming strategic actions (to try to act morally in each new situation encountered) by triggering instead fast, inexpensive, actions respecting, however, the same higher (moral or epistemic) norms as those involved in carefully controlled actions.[39]

[39] We noted, in Section 2.1.2, that fencers are trained to act impulsively rather than on the basis of their SAs. This is also a case of a critical use of fluency. The difference with the present case is that moral agents have a much wider scope of decisions open to them, in contexts that are also more variable.

Critical feeling is the basic concept that inspires such an attempt (Reber 2013). Critical feeling is "the strategic use of one's feelings in order to optimize outcome".[40] Aiming to exercise critical feeling through fluency, then, "means to know the consequences of fluency and to act accordingly".[41] The consequences of fluency, however, are Janus-faced. On the one hand, fluent actions—such as impulsive actions and habits—are felt as more pleasurable than effortful ones. On the other hand, fluent actions, being impulsive or habitual, are shortsighted: they tend to violate desirable epistemic or moral standards. The way out, then, is to turn morally and epistemically controlled actions into automatic, fluent ones. Systematic training in well-designed practices can shape the practioners' impulsive or habitual dispositions to act and to perceive the world in certain ways. As Reber and Slingerland (2011) have shown, Confucius designed a moral training based on the critical use of fluency, which is meant to allow trainees to acquire "a second nature" as moral agents. Confucius' project anticipates several recent findings in cognitive science. Confucius' first insight is that effortless action can only be attained by investing effort, namely through a strategic training designed to turn controlled into automatic decision-making.[42] His second insight is that investing effort generates its own reward: practice progressively raises the fluency of the corresponding action, which, in addition, is found to be coherent with the agent's moral values. A third insight is that fluent moral action can also be timely, i.e., that it can adjust to variable circumstances: "Timeliness" is the corresponding virtue cultivated by the sage.[43] Timeliness presupposes that one lets oneself be guided by one's own feelings—educated through prolonged practice. Hence, impulsive action is a key to timeliness, just as we found it to be in the cases of the expert chess player and of the fencer. The endeavor of self-cultivation can be seen as an attempt at reorganizing the structure of one's actions by consciously exploiting the constraints inherent to our primary trade-off.

4 Who is Operating the Switchboard?

4.1 A Subpersonal Operator?

Granting that agents may need to form their intentions to act on-line, how does the action switchboard operate? How, in other terms, is an agent determined to act in an impulsive, habitual, or strategic way? The assumption that the agent consciously selects the type of action it he intends to perform, along with selecting the content of this action, is incompatible with our time-resource trade-off. As indicated in Section 2.3.3.1, the subpersonal heuristic at work in determining whether to work on an habitual affordance or on a strategic goal has been claimed by neuroscientists to be based on the comparison of the respective prediction errors between the two systems for this particular action type. The system of acting that is expected to be the most reliable for

[40] Reber (2013), p. 174.
[41] Reber (2013), p. 175.
[42] Shiffrin and Schneider (1977).
[43] Slingerland (2001), p. 104.

achieving what the agent represents as his goal is the one that will automatically take the lead in controlling the present action.

Another powerful motive, discussed above, consists in the risk-reward ratio. In cases engaging fitness, an impulsive rather than a habitual action is automatically triggered. In cases engaging less urgent forms of utility, selection of an action type depends on the expected stakes, with lower stakes engaging habits, and higher stakes engaging SAs.

4.2 Self-Concepts and Action Preference

An additional dimension of selection should be mentioned: our types of action relate to the agents' own concept of self, and to the associated values. Habitual actions may be selected because they reflect, in the agents' own eyes and those of others, their own treasured distinctive abilities: they may choose to act, out of professional or domestic pride, in a way that best displays their acquired abilities. Other agents, in contrast, may take pride in their ability to plan new complex strategic actions, and enjoy discussing their plans with other agents. Self-related motivations, however, cannot trump the constraints of our time-resource trade-off. They can only bias them.

Although impulsive actions are not consciously selected, they can be consciously inhibited by agents able to adopt a critical view about their intentions to act. In contrast, some agents, like Harry Frankfurt's "wantons", are unable to inhibit their impulsions to act.[44] Alternatively, these may have, in the eyes of some agents, a special value. First, they involve no particular cognitive effort. Second, the induced dispositions to act feel retrospectively coherent with the agent's own present emotions and moods. In the case of mental impulsive actions, it has been shown that agents are more inclined to use cheap methods of assessment of their mental actions—based on fluency—when they feel happy and relaxed than when they are sad and depressed.[45] Third, being reactings, impulsive actions cannot be planned. Hence, their embodied expression cannot be easily faked.[46] This feature makes them a reliable predictor of the agent's trustworthiness. This may be why human communication is based in part on nonverbal, impulsive communicational gestures and intonations, and includes a significant number of impulsive utterances (such as interjections and expressive speech acts).

4.3 Impulsive Actions and Strategic Emotional Actions

In order to preserve their own social status, however, agents also frequently need to inhibit their impulsive dispositions to act. When sensing an affordance for expressing publicly their anger, for example, they can refrain from doing so. As noted by Frijda,[47] an impulsive action, when prevented from becoming overt, can still be produced covertly. The agent imagines, i.e. simulates the event of slapping his offender's face. Interestingly, an impulsive disposition to act can, then, be converted into planning a future action inspired by this feeling. For example, the agent can now intend to take revenge against the offender. Does this deferred action still qualify as an emotional

[44] Frankfurt (1988).
[45] Schwarz and Clore (2007).
[46] Ekman (1988).
[47] Frijda (1986), p. 191.

action? Pacherie (2002) claims that it does, and creates the category of a "semi-deliberate emotional action".[48] This type of action, according to Pacherie, is available to agents who are aware of the feeling that inspired their covert impulsive action and its associated tendency to act: they can, then, "convert" this feeling into the intention to act later in order to take revenge.

How does this proposal fare with our viewpoint on action from outside-in? One can fully agree with Pacherie that [intending to take revenge] prepares a planned, deliberate action. She takes it to be less than fully deliberate, however, because this intention is not integrated with the agent's other motivations. To this extent, it remains nearly as short-sighted as the initial, inhibited, impulsive action. She claims that there is, in addition, a category of fully deliberate emotional actions, based on truly reflexive emotional awareness of one's feelings. This category is based on the idea that one can make instrumental use of one's emotions. It is very similar to our notion of critical feeling-based strategic action, and thus, fully compatible with our proposal about strategic mental actions.

From the viewpoint defended here, however, taking some semi- and fully deliberate actions to be "emotional" does not make functional sense. First, impulsive and non impulsive types of action differ considerably in their phenomenology, their representational mode, their relation to time and the cognitive resources they engage. Second, how can an action tendency be "converted into" the conscious intention to pursue a certain goal? [49] An action tendency, on the present view, cannot survive a change in representational format and timing of action. Although there is a motivational coherence in the actions respectively involved in impulsive and "semi-deliberate" actions, embedding feelings in strategic actions deeply changes the rules of control precedence, and the other executive properties of the action. Thirdly, the fact that agents may fail to take advantage of the higher levels of the executive hierarchy described in Section 2.3 may suggest sub-typing SAs (according to the temporal depth of the constraints implemented), rather than cross-typing them (creating intermediate categories that are both emotional and strategic).

5 Conclusion

Naturalistic theories of action have been hampered by a persisting tension between an intellectualist conception of intention, taken to involve propositional attitudes with conceptual content, and the recognition of a variety of intentions, called "executive", "motor", "pragmatic", or "intentions in action": the latter, in contrast with "prior intentions", are not reflective, and, according to some theorists, are non-conceptual. We propose an alternative construal of the types of relation of intention to action.

A major constraint has shaped action systems in our phylogeny: Action must satisfy a trade-off between the time needed to complete an action, the amount of resources involved, and fitness significance. Assuming that no single type of action can adjust to incompatible requirements, there are at least three ways of enabling actions in cognitive systems, which apply to both bodily and cognitive actions: impulsive, routine, and strategic actions. We discussed how their respective representational formats optimize a given trade-off between timing and cognitive resources. The concept of intention as

[48] Pacherie (2002), p. 80.

[49] For a similar remark, associated with a different solution to this issue, see Scarantino (2014), p. 174.

classically understood may accordingly be distinguished into three types of causal-representational event: emotional affordance-sensing, utility affordance-sensing, and planning. The intentions to act impulsively or routinely are not detached from their affordance-sensings, in virtue of their being "pushmi-pullyu" appraisals. In common with the notion of "intention in action" our affordance-sensing (or, for short, "reacting"), is an intending simultaneous with an acting. Two differences, however, should be emphasized. First, the nonconceptual content of a reacting is structured by an appraisal of a given valence and intensity. Second the temporal patterns governing the onset and associated disposition to act of intentional reactings differ according to whether they are elicited by fitness-relevant circumstances or by recurrent utility. This temporal difference is determined by the constitutive urgency of risk or benefit in impulsive actions, and by the sense of a stable expectation of reward in habitual actions. Our last type of intention, strategic intention, corresponds to the classical concept of "prior intention". In contrast with the other two forms of intention, its content is conceptual, and depends on background propositional attitudes, namely beliefs, desires, and intentions. It is detached from its behavioral realization because planning, a strategic mental action present in every strategic physical action, precedes execution. It also controls and monitors it through a hierarchical structure of executive mechanisms organized by the temporal scales involved in planning.

Acknowledgments All my thanks to my colleague Dick Carter, for useful feedback on a prior version, and for his linguistic revision of the present article. I also thank Matias Baltazar, Laurence Conty, Terry Eskenazy, Martin Fortier, John Michael and two anonymous reviewers for their comments and/or bibliographical suggestions. This research has been supported by an ERC Advanced Grant "Dividnorm" # 269616, and by two institutional grants: ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL.

References

Balcomb, F.K., and L. Gerken. 2008. Three-year old children can access their own memory to guide responses on a visual matching task. *Developmental Science* 11(5): 750–760.

Bar-On, D. 2013. Origins of meaning: Must we go Gricean? *Mind & Language* 28(3): 342–375.

Barrett, L.F., and M. Bar. 2009. See it with feeling: Affective predictions during object perception. *Philosophical Transactions of the Royal Society, B: Biological Sciences* 364(1521): 1325–1334.

Bernard, S., Proust, J., and F. Clément (under review). The medium helps the message: Early sensitivity to auditory fluency in children's endorsement of statements. *Frontiers in Psychology.*

Bratman, M.E. 1987. *Intention, plans, and practical reason*. Cambridge: Harvard University Press.

Burgess, P.W., A. Quayle, and C.D. Frith. 2001. Brain regions involved in prospective memory as determined by positron emission tomography. *Neuropsychologia* 39(6): 545–555.

Carruthers, P., and J. B. Ritchie. 2012. The emergence of metacognition: Affect and uncertainty in animals. In *Foundations of metacognition*, eds. M. Beran, J. Brandl, J. Perner, J. Proust, 76–93.

Couchman, J.J., M.J. Beran, M.V.C. Coutinho, J. Boomer, and J.D. Smith. 2012. Evidence for animal metaminds. In *The foundations of metacognition*, ed. M.J. Beran, J. Brandl, J. Perner, and J. Proust, 21–35. Oxford: Oxford University Press.

Damasio, H., T. Grabowski, R. Frank, A.M. Galaburda, and A.R. Damasio. 1994. The return of Phineas Gage: Clues about the brain from the skull of a famous patient. *Science* 264(5162): 1102–1105.

Dreyfus, H., and S.D. Kelly. 2007. Heterophenomenology: Heavy-handed sleight-of-hand. *Phenomenology and the Cognitive Sciences* 6(1–2): 45–55.

Ekman, P. 1988. Lying and nonverbal behavior: Theoretical issues and new findings. *Journal of Nonverbal Behavior* 12(3): 163–175.

Fleming, S.M., J. Huijgen, and R.J. Dolan. 2012. Prefrontal contributions to metacognition in perceptual decision making. *The Journal of Neuroscience* 32(18): 6117–6125.
Frankfurt, H. 1988. *The importance of what we care about*. Cambridge: Cambridge University Press.
Frijda, N.H. 1986. *The emotions*. Cambridge: Cambridge University Press.
Fuster, J.M. 1989. *The prefrontal cortex*. New York: Raven.
Gibson, J.J. 1977. The concept of affordances. In *Perceiving, acting and knowing*, ed. R. Shaw and J. Bransford, 67–82. Hillsdale: Erlbaum.
Grafman, J. 2002. The structured event complex and the human prefrontal cortex. In *Principles of frontal lobe function*, ed. D.T. Stuss and R.T. Knight, 292–310. Oxford: Oxford University Press.
Greifeneder, R., and C. Unkelbach. 2013. Experiencing thinking. In *The experience of thinking*, ed. C. Unkelbach and R. Greifeneder, 1–8. Hove: Psychology Press.
Hursthouse, R. 1991. Arational actions. *Journal of Philosophy* 88(2): 57–68.
Kepecs, A., U. Naoshige, H. Zariwata, and Z.F. Mainen. 2008. Neural correlates, computation and behavioural impact of decision confidence. *Nature* 455: 227–231.
Koechlin, E., C. Ody, and F. Kouneiher. 2003. The architecture of cognitive control in the human prefrontal cortex. *Science* 302(5648): 1181–1185.
Koriat, A. 2000. The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition* 9: 149–171.
Lee, S.W., S. Shimojo, and J.P. O'Doherty. 2014. Neural computations underlying arbitration between model-based and model-free learning. *Neuron* 81(3): 687–699.
McDaniel, M.A., and G.O. Einstein. 2000. Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology* 14(7): S127–S144.
Mele, A.R. 1992. *Springs of action: Understanding intentional behavior*. Oxford: Oxford University Press.
Millikan, R.G. 1995. Pushmi-pullyu representations. *Philosophical Perspectives* 9: 185–200.
Niv, Y., D. Joel, and P. Dayan. 2006. A normative perspective on motivation. *Trends in Cognitive Sciences* 10(8): 375–381.
Ouellette, J.A., and W. Wood. 1998. Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin* 124(1): 54.
Pacherie, E. 2002. The role of emotions in the explanation of action. *European Review of Philosophy* 5: 55–90.
Pacherie, E. 2006. Toward a dynamic theory of intentions. In *Does consciousness cause behaviour? An investigation of the nature of volition*, ed. S. Pockett, W.P. Banks, and S. Gallagher, 145–167. Cambridge: MIT Press.
Pacherie, E. 2011. Non-conceptual representations for action and the limits of intentional control. *Social Psychology* 42(1): 67–73.
Prinz, J.J. 2004. *Gut reactions: A perceptual theory of emotion*. Oxford: Oxford University Press.
Proust, J. 2003. How voluntary are minimal actions? In *Voluntary action*, ed. S. Maasen, W. Prinz, and J. Roth, 202–221. Oxford: Oxford University Press.
Proust, J. 2013. *The philosophy of metacognition. Mental agency and self-awareness*. Oxford: Oxford University Press.
Proust, J. 2014a. What are feelings about? To appear on line. In *Open MIND*, ed. T. Metzinger and J.M. Windt. Frankfurt: MIND Group.
Proust, J. 2014. Epistemic action, extended knowledge, and metacognition. *Philosophical Issues.*
Reber, R. 2013. Critical feeling. In *The experience of thinking*, ed. C. Unkelbach and R. Greifeneder, 173–189. Hove: Psychology Press.
Reber, R., and E.G. Slingerland. 2011. Confucius meets cognition: New answers to old questions. *Religion, Brain & Behavior* 1(2): 135–145.
Scarantino, A. 2014. The motivational theory of emotions. In *Moral psychology and human agency*, ed. D. Jacobson and J. D'Arms. Oxford: Oxford University Press.
Schwarz, N., and G.L. Clore. 2007. Feelings and phenomenal experiences. In *Social psychology: Handbook of basic principles*, ed. A.W. Kruglanski and E.T. Higgins, 385–407. New York: Guilford Press.
Searle, J.R. 1983. *Intentionality, an essay in the philosophy of mind*. Cambridge: Cambridge University Press.
Sellars, W.S. 1973. Actions and events. *Noûs* 7: 179–202.
Shiffrin, R.M., and W. Schneider. 1977. Controlled and automatic human information processing, II. Perceptual learning, automatic attending, and a general theory. *Psychological Review* 84(2): 127–189.
Slingerland, E. 2001. Virtue ethics, the analects, and the problem of commensurability. *Journal of Religious Ethics* 29(1): 97–125.
Smith, R.E. 2003. The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29(3): 347–361. doi:10.1037/0278-7393.29.3.347.