Control: conscious and otherwise

Article *in* Trends in Cognitive Sciences · August 2009

Impact Factor: 21.97 · DOI: 10.1016/j.tics.2009.04.010 · Source: PubMed

Pat Churchland
University of California, San Diego

Cell PRESS

Control: conscious and otherwise

Christopher L. Suhler and Patricia S. Churchland

Department of Philosophy, University of California, San Diego, La Jolla, CA 92093, USA

Social psychologists have shown human decisions to be sensitive to numerous ordinary, possibly nonconscious, situational contingencies, motivating the view that control is largely illusory, and that our choices are largely governed by such external contingencies. Against this view is evidence that self-control and goal-maintenance are regularly displayed by humans and other animals, and evidence concerning neurobiological processes that support such control. Evolutionarily speaking, animals with a robust capacity to exercise control – both conscious and nonconscious – probably enjoyed a selective advantage. Counterbalancing data thus point to an account of control that sees an important role for nonconscious control in action and goal maintenance. We propose a conceptual model of control that encompasses such nonconscious control and links in-control behavior to neurobiological parameters.

Introduction

An important notion in moral philosophy and many legal systems is that certain circumstances can mitigate an individual's responsibility for a transgression. Generally speaking, such situations are considered extenuating in virtue of their exceptional influence on a person's ability to act and make decisions in a normal manner. The essence of the case for diminished responsibility is that these special circumstances impede the ability of a normal person to exercise self-control. Although differences exist between jurisdictions, the basic point is that if the circumstances put exceptional pressure on the capacity for control, the judge or jury can reduce the penalty.

In recent years, however, this notion of diminished responsibility has come to wider attention in a quite unexpected way. Some researchers, drawing on findings from social psychology, have argued that situational forces could have a much larger role in behavior than traditionally assumed. The situational forces in question are often entirely ordinary, mundane and seemingly trivial. Given that such influences are pervasive, the general issue raised concerns control in commonplace cases. According to a condensed version of this view – which for convenience we shall call the 'Frail Control' hypothesis – even in unexceptional conditions, humans have little control over their behavior. If correct, this line of argument could have widespread and dramatic ramifications, notably for our practices of attributing moral and legal responsibility.

Before proceeding, we note that although control is usually linked to responsibility, control does not always entail full responsibility, as for example when a seriously delusional subject lies in wait with a rifle to kill the werewolf disguised as a bus driver. Conversely, someone might kill while under the influence of alcohol, but will be held responsible nonetheless. Typically, however, a tight connection exists between the two, both legally and informally. Thus, even though in certain rare cases control and responsibility come apart, in most cases of moral and legal responsibility attribution, control and responsibility are closely linked. Our discussion is relevant mainly to such cases.

Although we agree that moral philosophy and the law can benefit from a greater understanding of developments in psychology and neuroscience, we suggest that the Frail Control challenge is markedly weakened once a wider range of data is considered. In our assessment, the Frail Control hypothesis underestimates the vigor of normal goal-maintenance in the face of distractions, and neglects the role of nonconscious aspects of control as displayed, for instance, in the exercise of cognitive, motor and social skills.

The Frail Control hypothesis might also reflect contagion from the demise of the noncausal theory of free will [1,2]. Although the Cartesian idea of 'uncaused choice' probably is illusory, executive control, by contrast, seems entirely real – even when exercised nonconsciously. But whereas conscious control (or its absence) is a notion that most people readily grasp, it is less obvious exactly how nonconscious control should be understood. A large psychological literature has demonstrated that nonconscious, automatic processes are pervasive and anything but 'dumb'. Instead, they are often remarkably sophisticated and flexible in performing functions such as goal pursuit that were once considered the sole province of conscious cognition [3]. Even so, this body of findings has rarely been explicitly connected to the issue of control broadly construed. (This is, however, beginning to change; see, for instance Refs [4–6]).

Here, we develop an account of control that we believe goes some way towards sharpening the meaning of control – including nonconscious control – in a way that accommodates the role of nonconscious processes in nearly everything we do. The general conclusion we will be arguing for – that nonconscious processes can support a robust form of control and, by extension, that consciousness is not a necessary condition for control – is one that builds substantially on prior research and which others [4,5] are also working to establish. The notable contribution of our account to this project is a model of control in which neurobiological criteria, rather than intuitive or behavioral criteria alone, define the boundaries of control. A notable virtue of this account, in light of the pervasiveness of automatic processes in our cognitive lives, is that it is agnostic as to whether the underlying processes are

Corresponding author: Suhler, C.L. (csuhler@ucsd.edu).

1364-6613/$ – see front matter doi:10.1016/j.tics.2009.04.010

conscious or nonconscious. Before developing this perspective on control, however, we will describe in a bit more detail the contours of the Frail Control hypothesis and its manifestations in the philosophical and psychological literatures.

Frail control

A leading advocate of the Frail Control hypothesis is the philosopher John Doris (e.g. Ref. [7]). He bases his claims on a range of data from social psychology showing that choices can be affected by various manipulations, such as priming (often below the level of consciousness) or ostensibly banal environmental features. For example, subjects exposed to words related to rudeness on a scrambled-sentence task are subsequently more likely to interrupt a (staged) conversation between the experimenter and another person than are subjects primed with words related to politeness or controls who are not primed [8]. Other studies show that people are more likely to litter in a particular setting when it is heavily littered than when the same setting is clean [9] (for reviews, see Refs [3,10]).

Adding to the surprise, the data seem to show that very minor environmental influences can at times produce large effects, perhaps overwhelming long-term goals and character traits. Among the examples Doris cites are the finding by Isen and Levin [11] that '[p]assersby who had just found a dime were twenty-two times more likely to help a woman who had dropped some papers than passersby who did not find a dime' ([12], p. 34) and the finding by Darley and Batson [13] that '[p]assersby not in a hurry were six times more likely to help an unfortunate who appeared to be in significant distress than were passersby in a hurry' ([12], p. 34). More recently, in the study on littering behavior noted earlier, 33% of subjects littered in the 'order' condition (in which there was no graffiti on the walls of an alleyway), whereas in the 'disorder' condition (graffiti on the walls) that percentage more than doubled, to 69% [9]. Although most effects in social psychological studies are not this dramatic, findings such as these nevertheless demonstrate an extreme of environmental influence on behavior that is sufficient to render the Frail Control hypothesis a live possibility.

These data are connected to the issue of responsibility in the following way: if your choice is strongly affected by situational factors in ways that you are unaware of, then you plausibly have an excuse for your actions. Hence, taking a prohibited shortcut through a field in an area with lots of graffiti, garbage and wrecked bicycles is not really your fault. Unbeknownst to you, your brain responds to highly disordered situations by relaxing your usual inhibitions against violating a No Trespassing sign. You had little or no control over that effect. Doris ([14], p. 136) echoes widespread philosophical assumptions when he says that to be responsible we must have 'normative competence', meaning that we consciously weigh the evidence, effectively deliberate, and make a decision. If the deciding and weighing is below the level of consciousness, normative competence is compromised. No normative competence, no responsibility.

The appeal of the Frail Control hypothesis is not limited to Doris or even to philosophers. For instance, Timothy Wilson [15], in describing the implications of what he calls the 'adaptive unconscious', says that '[w]e may have the impression that we, our conscious selves, are in complete control, but that is at least in part an illusion' ([15], p. 48). See also Harman [16], Bargh [17] (who, we note, should not be considered a true proponent of Frail Control given his research explicitly connecting nonconscious processes to control [4,5]), Wegner [18] and Appiah [19], along with a recent news feature in *Nature* [20].

The conclusion that our actions are much more frequently excusable than hitherto assumed could have monumental implications for the law, both criminal and civil, and our daily social interactions. A rather different picture of control emerges, however, once the range of data is expanded to include neurobiological, clinical and other behavioral data, together with considerations from evolutionary biology. Broadening the range of data rebalances the inquiry, with the result that control seems to be a more robust, but also more complex, phenomenon than envisioned by the Frail Control hypothesis. We will begin by discussing the relevant evolutionary data, which provide the ultimate adaptive rationale for the existence of the neurobiological mechanisms supporting control functions to be discussed later.

The co-evolution of control and situational responsiveness

The co-evolution of sensitivity in responding to a diverse array of environmental stimuli and the capacity for executive control is highly probable [2,21]. Generally speaking, if an organism is to reap the benefits of adaptive responsiveness to its environment, it must also be able to control how and to what it responds. Even when some responses are strongly reflexive, such as crying out when suddenly experiencing pain, they must also be modifiable if the animal needs to suppress the response.

Observations of mammalian behavior suggest that mature animals do indeed regularly exhibit control. A cougar that can carefully stalk a deer will do better than one who just runs after it; antelope that go skittering off every time they glimpse a lion in the distance are apt to waste excessive amounts of energy. Similarly, in hunter-gatherer conditions, a human hunter who plans, takes advantage of past knowledge, waits when necessary and so forth is more likely to get large game than an impulsive hunter who might have to settle for scrawny rabbits or nothing at all. Additionally, laboratory experiments show that rats can defer gratification if doing so allows them to get substantially more food than inflexibly going for a quick but smaller reward [22]. Rats can also be trained to stop a bar press when a tone is signaled, even if the movement has already been initiated [23]. In the rats, relevant behavioral differences in these tasks are described as differences in the capacity for control, and the circuitry underlying these capacities is an object of study.

Mechanisms for exercising control in numerous species, hominins included, were probably selected for in conditions that favored being able to defer gratification, wait for the advantage, plan ahead, undertake a complex, multi-step action and so on. In discussing control as a deep and general feature of animal behavior, Baumeister [21] makes the point that the desire for control, both of physical and social conditions, is fundamental to reproductive success. He thus remarks that 'if control is part and parcel of getting most of the things one wants in life, a person could evade wanting control only by not wanting anything' ([21], p. 96). The pursuit of goals and achievement of them requires some measure of control, and the longer the lag time or the more obstacles in the path, the greater the need for control. (Of course, some species have greater capacities for planning and following through than others, and individual variation within a species is seen as well.)

In the social environment of one's own species, the capacity to exercise control and select an appropriate action is perhaps even more crucial. For example, hierarchy is extremely important in chimpanzee troops, and young chimpanzees must learn to control impulses such as the urge to take food when a higher-ranking chimpanzee also sees the food [24]. Low-ranking baboon males, however lusty, are very cautious about mating with a female unless they can do so undetected by the alpha male [25]. And a young male baboon seeking entry into a new troop must bide his time, carefully determine which high-ranking female he might successfully associate with, make himself agreeable to her, and slowly win acceptance. The brash outsider who tries to bypass these social niceties by forcing himself on the troop is shunned, and fares poorly [26].

In our modern culture, exercising control to adjust to and thrive in one's social environment is likewise paramount. In line with this, Baumeister [21] observes that humans have an expanded repertoire of ways to satisfy the desire for control. Humans exhibit control by, for instance, attending school, learning to build a house, maintaining a garden or farm animals, or going to work regularly, and such control tends to pay off over the course of a lifetime [27,28]. Millions of people get out of a warm, cozy bed in the morning to go to a job they do not much like – a rather clear demonstration of some measure of quotidian control. Most humans are also acutely aware of the reputational ramifications of behavior, and regularly suppress their desires to insult, fisticuff, cheat, grope and so forth when it is likely to do damage to their reputation. Skills of self-discipline and self-control are acquired by maturing children as a result of social pressure from many directions, including from peers [29]. The dopamine system is known to play a very important part in this process of developing normal social and cognitive abilities, including the internalization of social skills [30].

In circumstances where nothing much hangs on doing A rather than B, vigilance might be lower and situational factors more influential. When pursuing a goal, one can encounter many 'fringe' choices – whether or not to pick up a piece of litter, for example. Nevertheless, how one decides these fringe choices has very little to do with the normal function of executive control in pursuit of a goal. Suppose you are working on an important grant with a looming deadline, but a colleague asks for your help moving a piece of furniture into his office. While attending to the interrupting task, people typically experience frequent intrusive thoughts about the goal, getting back to the goal, how to complete the current task quickly, and so on. Demonstrated experimentally and sometimes referred to as the 'Zeigarnik effect' [31,32], this phenomenon implies that nonconscious processes continue to keep the goal high in priority until resumption of the goal-related action, no matter the interruption by task-irrelevant contingencies. Rather than frail control, this phenomenon and the others described earlier bespeak rather stalwart and sturdy control, although of course individual variability must always be acknowledged. A natural question given the pervasive behavioral manifestations of control is this: what are the neurobiological mechanisms that support such control (Box 1 and Figure 1)?

Box 1. What can neuroscience tell us about being 'in control'?

A wide range of neurobiological findings – both anatomical and physiological – provide insight into how control is maintained, and how it can be compromised. Although much remains to be discovered, prefrontal and limbic structures have long been implicated in control, and research has focused on trying to identify more precisely the neurobiological mechanisms underlying control. Studies of clinical subjects by Damasio and colleagues [38–40] have revealed the particular importance of the orbital and ventral medial regions of the prefrontal cortex in normal (i.e. non-pathological) executive control, a view supported by many studies [37]. Research by Aron and Poldrack using Go-Stop paradigms (e.g. Refs [59,60]) has identified three connected areas as important players in the regulation of control and inhibition: the right inferior frontal cortex, the pre-supplementary motor area (pre-SMA) and the basal ganglia, notably the subthalamic nucleus [33] (Figure 1a in main text). Tensor diffusion imaging shows these areas in humans to be highly interconnected (Figure 1b in main text). At the molecular level, the role of various neurochemicals, along with their relative balance, is a crucial part of the story. For example, low serotonin levels are associated with histories of impulsive, violent behavior in humans [42–44] and in rodents and non-human primates [46]. Studies on addiction are revealing the role of dopamine, opioid peptides and corticotrophin releasing factor (CRF) in control functions [41,42].

Moreover, theoretical models have begun to provide a conceptual framework within which to understand how the properties of neural networks yield the functional process we call executive control. Miller and Cohen [34] have developed a model that describes prefrontal cortex function as a form of 'top-down' control that guides behavior by activating internal representations such as goals and intentions. They propose, in essence, that the prefrontal cortex exercises control by maintaining goal-related patterns of activity and sending bias signals that modulate activity in other areas of the brain in ways that promote the performance of goal-relevant behaviors [34].

A neurobiological account of control

We suggest that a range of neurobiological data and models of brain function (such as those described in Box 1) point to a way to sharpen the meaning of 'control'. Our proposal has two parts. The first component is anatomical, specifying that the brain regions and pathways implicated in control are intact and that behavior is regulated by these mechanisms in a way consistent with prototypical cases of good control. So, for instance, if trauma or disease damages areas implicated in control – such as the fronto-basal-ganglia circuit [33] and prefrontal cortex [34] – control will be impaired [35–40]. The second component is physiological, and includes the molecular mechanisms whereby control is regulated. Even if the anatomical structures for control functions are intact, functionality requires that

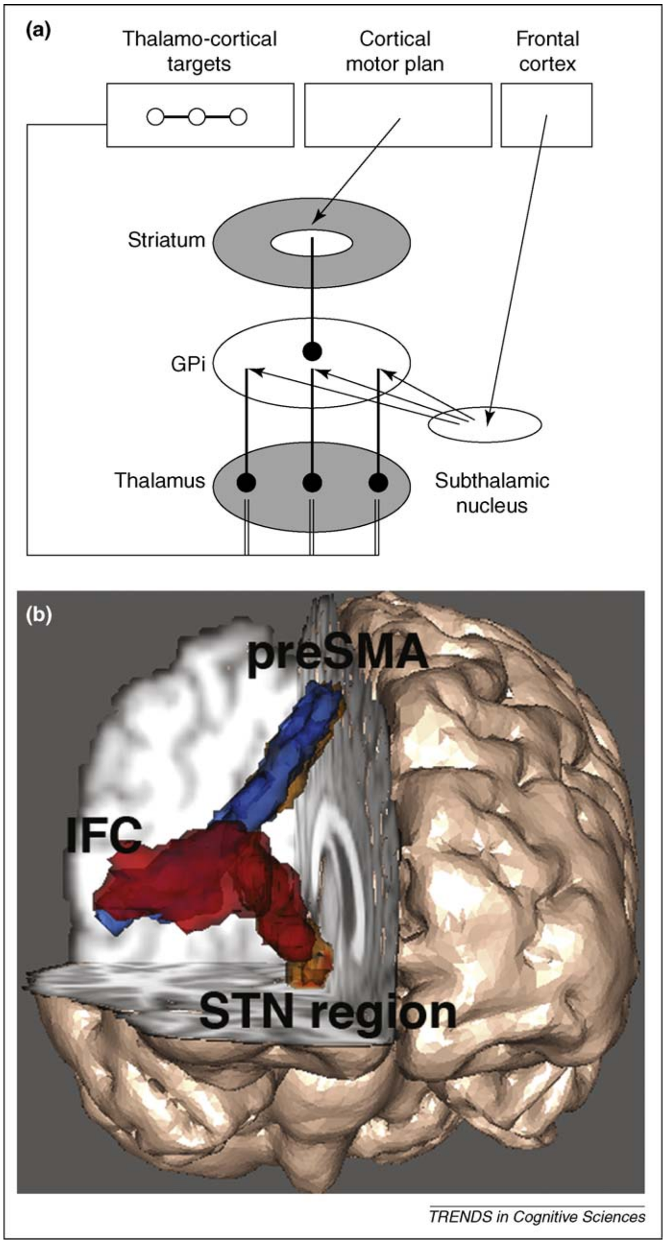

Figure 1. The fronto-basal-ganglia circuit. **(a)** Schematic of the fronto-basal-ganglia network for 'going' and 'stopping', which has been implicated as an area supporting inhibitory control of action and thought. The 'go' process is generated by the premotor cortex, which excites the striatum and inhibits the globus pallidus, removing inhibition from the thalamus and exciting the motor cortex. The 'stop' process could be generated by the inferior frontal cortex, leading to activation of the subthalamic nucleus, which, in turn, increases excitation of the pallidum and inhibits thalamocortical output, reducing activation in the motor cortex. **(b)** Diffusion-weighted imaging of white matter tracts connecting the dorsomedial pre-supplementary motor area (preSMA), the inferior frontal cortex (IFC) and the subthalamic nucleus (STN region). This provides anatomical backing – albeit tentative because the results have not yet been replicated in postmortem studies – for the proposed circuit depicted in (a). Figures and descriptions courtesy of Adam Aron. Reproduced, with permission, from Ref. [33].

the levels of various neurochemicals – neurotransmitters, hormones, enzymes and so on – are maintained normally. To a first approximation, what 'normal' means here will be determined experimentally by discovering links between uncontentious examples of control in behavior, and the neurobiological parameters in question; similarly for cases of impaired control. Roughly speaking, and granting individual variability, the normal range of the implicated neurochemicals (for a given species) is calibrated to the spectrum of values that the brain evolved to maintain in response to environmental demands typical of the species' evolutionary past. Outside this range, control will be compromised. For instance, in addicts, a delayed return to baseline of corticotrophin releasing factor (CRF) levels, correlated with high levels of anxiety, seems to be an important factor in addicts' recidivism [41,42]. Considering a different parameter, low serotonin levels are correlated with poor impulse inhibition, implying that this neurochemical has an important role in control [43–46].

Is there a way to connect this neurobiological perspective on what constitutes being 'out of control' with the sorts of 'situational' factors cited in support of the Frail Control hypothesis? We suggest not. By contrast, the data can readily be connected with the unusual extenuating circumstances adduced in certain criminal cases. It is well-known that the normal levels of neurochemicals, and thus control, can be disrupted when external circumstances are, for instance, profoundly threatening. Great fear or shock can trigger a cascade of stress responses (including a rise in CRF, glucocorticoids and the catecholamines epinephrine and norepinephrine [41,42,47,48]) that could cripple control mechanisms. As a result, a man about to be executed might wet his pants, a captured spy might divulge secrets after 'being shown the instruments of torture' or a cuckolded husband might knife the disgraced pair in bed. These are the kinds of circumstances that courts regularly consider when asked to reduce penalties. Crucially, however, they are not the kind of circumstances on which the Frail Control hypothesis relies.

To be clear, the account just sketched does not set the unreasonable standard that every relevant neurochemical must be at its ideal level or even within its normal range. Instead, the physiological requirement for being in control is defined in terms of a hyper-region in an n-dimensional 'control space'. An important consequence of defining control in this way is that there will be many different combinations of neurochemical levels that fall within the 'in control' hyper-region. As a result, a given neurochemical straying outside its normal range need not render a person 'out of control', assuming that other neurochemicals are within their normal ranges and that the deviation is not too extreme (see Figure 2 for a rough depiction of how this might look in three dimensions).

The role of nonconscious processes

In the examples discussed towards the end of the previous section, the relevant situational factors (e.g. imminent execution) are ones of which the individual is conscious and which are being managed (at least to some substantial extent) by conscious cognition. But what does the aforementioned account have to say about whether nonconscious processes are capable of supporting genuine control?

According to a traditional framework – which, begging some forbearance, we will call 'neo-Kantian' – consciousness is a paramount, and perhaps even necessary, condition for a decision's being considered free. According to the neo-Kantian, consciousness must play a substantial part in most or all steps leading to a free decision: deliberating, choosing, intending and acting. The interplay of

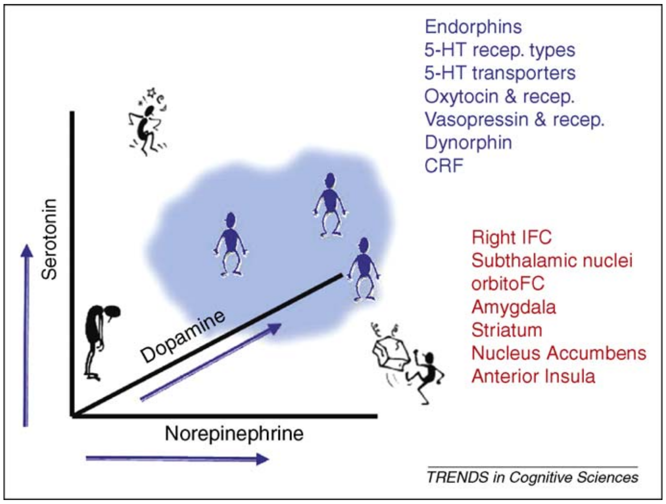

Figure 2. A 3D control space. A cartoon of what a control space defined only in terms of three important neurochemicals – dopamine, serotonin and norepinephrine – might look like. In reality, a control space would have many parameters, including also for example, density of serotonin receptor types, serotonin transporter values, levels of dynorphin and corticotrophin releasing factor (CRF), in addition to values expressing connectivity between brain structures such as the ventromedial frontal cortex and amygdala, and between the dorsolateral prefrontal cortex and anterior cingulate cortex. The blue cloud represents the 'in control' region. Many different combinations of values of the various parameters will fall within this region and thus be correlated with behavior that is 'in control', This captures the likelihood that there are different ways of being in control and certainly different ways of being out of control. The boundaries of the space are somewhat blurry, representing control as graded rather than an all-or-nothing matter. Regrettably, the diagram cannot provide for dynamical properties that would reflect changes over long and short periods of time. Abbreviations: 5-HT, serotonin; CRF, corticotrophin releasing factor; right IFC, right inferior frontal cortex; orbitoFC, orbitofrontal cortex.

reasons in deliberation must be transparent because a reason must be conscious to be a reason at all; otherwise, it is a mere cause. Control, accordingly, is believed to be limited to those cases where most or all evidence, reasons, weighting of reasons and so forth that contribute to a choice are consciously accessible. This transparency is central to the emphasis placed on consciously deliberated choice as a paradigmatic case of control in much of the philosophical literature, and to the relation between 'normative competence' and responsibility invoked by Doris ([14], p. 136). In keeping with the neo-Kantian perspective, the Frail Control hypothesis implicitly attaches enormous importance to whether the factors that have a role in an action can be consciously acknowledged as reasons. In our view, however, the general consciousness requirement for being 'in control' is unrealistic. Exactly what role awareness of specific factors must have for an action to be considered controlled, relative to neurobiological criteria, is a matter not of stipulation, intuition or semantics, but scientific discovery.

The exercise of skills is one domain where nonconscious processes are entirely consistent with – and even boost – successful control. Skilled responses are involved in just about everything we do, from driving, reading and gardening, to getting along with members of our community and finding our way home [49]. What has emerged from studies of automatization of motor skills is that all parts of the motor network reduce their levels of activity as the skill is acquired [50]. A recent finding also shows that the parts of the network (the cerebellum, cingulate, motor area, supplementary motor area and putamen) have stronger effective connectivity, and thus influence each other more, after automaticity is achieved [51]. (Also worth noting is that not all automatic behaviors are the same. For example, pupillary dilation and contraction is automatic, and not something over which we have control. By contrast, the automatic behaviors displayed by violinists and basketball players are highly controlled.)

Studies of skill acquisition, whether motor [52] or cognitive [53], also indicate that in skilled or trained individuals, conscious attention is directed not to the intermediate steps, but to the larger aim and to unforeseen hazardous contingencies. Routine control can therefore be automatic (as evidenced, for instance, by increases in anterior cingulate cortex activity with practice [53]), while vigilant control can be directed to other things. A firmly held goal often means that potential distractions are nonconsciously ignored, and that disruptive emotions or drives are nonconsciously suppressed. When social niceties become 'second nature', one does not have to consciously work out what to do, or consciously suppress intentions that could intrude and make for awkwardness. Unlike children, adults do not have to consciously remind themselves not to break wind in polite company or to shake hands upon meeting someone. Thus, habit and routine serve to spare the brain the energetic costs of close attention and to give the benefits of smooth operation [6,54–56], making nonconscious control of this sort a great energy- and face-saving device. Notably, cognitive, motor and social skills, including those that underlie habit and routine, are often invoked in later explanations of actions and are certainly robust enough in their guidance of action to be considered genuine reasons.

Additional support for the value of automaticity comes from the hypothesis that (conscious) executive control is itself a somewhat limited resource. According to this view, known as the 'self-regulatory resource model', the amount of energy people have to expend on conscious self-regulation is limited, with the result that expending it on one task reduces the amount available for other tasks [57,58]. This suggests that nonconscious processes not only perform control functions of their own but can also help to ensure the efficacy of conscious mechanisms of control.

Furthermore, as evidenced by our ability to function while being bombarded by stimuli on a moment-to-moment basis, environmental factors – even if processed below the level of conscious awareness – do not flow straight through to trigger behavior. A reason for this is suggested by the model developed by Miller and Cohen [34], which, as sketched in Box 1, proposes that the prefrontal cortex exerts control by sending bias signals that modulate activity in other brain areas. On this account, the totality of environmental influences – via automatic processes – clearly need not determine behavior. Even if, acting alone, environmental factors were to give rise to a pattern of activity divergent from a goal, the prefrontal cortex could, through bias signals, cause a goal-relevant pattern of activity to prevail instead.

In summary, although the idea that reasons and control can be (and often are) nonconscious is unacceptable to those who – even tacitly – accept the traditional, neo-Kantian view of action, it is consistent with the data. We return to the connection between conscious and nonconscious control in the penultimate section of this article. Before that, however, we would like to return to, and attempt to answer in light of the preceding discussion, the question about the implications of social psychological findings posed in the section on Frail Control.

Social psychology and control

The implications of the neurobiological account of control developed here for the interpretation of social psychological results can be summarized as follows: quite simply, most of the patterns of behavior described in the social psychology literature do not fall outside the realm of control (see also Ref. [3]). The reason is that although the effects studied by social psychologists are mediated by situational factors and (often) by nonconscious processes, evidence indicates that the requirements for control set out earlier are typically met. First, the brain structures essential for control functions are intact (the anatomical condition). And second, the circumstances studied are usually within the typical range encountered in the evolutionary past of humans, and thus, levels of various neurochemicals on which the proper functioning of the anatomical structures depends can reasonably be expected to be within their normal ranges (the physiological condition). Contrary to the claims of the Frail Control hypothesis, therefore, findings from social psychology – brilliant although they might be – should not be taken to motivate a substantial revision of our moral and legal practices of responsibility attribution.

The relationship between conscious and nonconscious control

To be clear, we are not advancing the radical thesis that there is no such thing as consciousness or conscious control. The main point of this article is rather that although consciousness – for instance of goals and what the neo-Kantian would call 'reasons' – does sometimes have an important role in control, it is not required for control. Nonconscious control can be – and frequently is – exercised, and this control can be every bit as genuine as the conscious variety.

Given that the notion of nonconscious control is only beginning to gain traction in the scientific literature [6], we are not currently in a position to speculate about the precise interplay of conscious and nonconscious processes in controlled behavior or about the similarities and differences between the neurobiological substrates for these processes. Even so, it seems a safe bet that anatomical and physiological factors such as those discussed will figure in both conscious and nonconscious control in some fashion, and to some degree. The real work, however, will be in investigating the neurobiological details and teasing out how the anatomical and physiological factors underlying conscious and nonconscious control coincide and differ.

Moreover, although the extremes of conscious and nonconscious control might be fairly clear, the gradations and connections between cases are not yet central targets of research, let alone known with any certainty. Indeed, the precise roles of consciousness in brain function, along with its mechanisms, are not well-understood. One appealing, but admittedly speculative, idea is that understanding the relationship between conscious and nonconscious control could provide new and more penetrating insight into these broader questions about consciousness. By examining a range of cases between the extremes of conscious and nonconscious control and systematically varying the situational parameters, researchers might be able to illuminate the factors that prompt increased conscious involvement in the process of decision and action. This question is just one of many on which much work remains to be done (Box 2).

Box 2. Questions for future research

- To what extent do the neurobiological mechanisms subserving conscious and nonconscious control overlap? To what extent are they different, and if substantial differences exist, what specific anatomical and physiological factors mediate these differences?
- Relatedly, what task demands and experiential factors drive the transition from nonconscious to conscious control (and vice versa)? Can greater knowledge of what factors contribute to control being conscious (or not) provide insight into the more general function of consciousness in cognition?
- How do manipulations such as those employed in the social psychology literature affect the neurobiological variables described in the text?
- What brain areas, in addition to those described in the text, have large roles in control (especially nonconscious control)?
- To what extent are there robust species and individual differences in nonconscious control? What anatomical and physiological differences contribute to such variation? What changes occur in the brain when children are actively taught to exercise self-control over their emotions, impulses and desires?
- How do learning and practice affect the functioning of nonconscious control mechanisms, and what neurobiological alterations underlie these changes?

Conclusion

Recent challenges to the classical framework of control and responsibility are based on data from social psychology showing that minor external contingencies can have a substantial role in behavior even when we are unaware that they do so. The data are taken to imply that control is rare and frail, and that the category of excuses from moral and legal responsibility should be modified accordingly. From our perspective, once these findings are placed alongside a broader range of data, a very different hypothesis is motivated: goal-maintenance and executive control are remarkably robust, and elements of control are often nonconscious. Neurobiological grounding for this alternative hypothesis is provided by findings that suggest a framework for control based on anatomical and physiological parameters. Accordingly, it is possible to model control as neural activity within a parameter space, where a region of the space characterizes values of various neurobiological parameters needed for executive control. So long as control-relevant anatomical structures are intact and the neurochemicals on which their functionality depends are within their appropriate ranges, sensitivity to situational contingencies and nonconscious processes are appropriate aspects of control and goal-directed behavior, not obstacles to them.

Acknowledgements

We thank Paul Churchland, John Doris and anonymous reviewers for helpful comments on an earlier draft of this paper.

References

1 Churchland, P.S. (2002) *Brain-Wise: Studies in Neurophilosophy*. MIT Press

2 Dennett, D.C. (2003) *Freedom Evolves.* Penguin
3 Bargh, J.A. and Morsella, E. Unconscious behavioral guidance systems. In Then a Miracle Occurs: Focusing on Behavior in Social Psychological Theory and Research (Agnew, C. et al., eds.), Oxford University Press (in press)
4 Bargh, J.A. and Huang, J.Y. (2009) The selfish goal. In *The Psychology of Goals* (Moskowitz, G. and Grant, H., eds), pp. 127–150, Guilford
5 Hassin, R.R. *et al.* (2009) Automatic and flexible: The case of nonconscious goal pursuit. *Soc. Cogn.* 27, 20–36
6 Berkman, E.T. and Lieberman, M.D. (2009) The neuroscience of goal pursuit: Bridging gaps between theory and data. In *The Psychology of Goals* (Moskowitz G., G. and Grant, H., eds), pp. 98–126, Guilford
7 Doris, J.M. (1998) Persons, situations, and virtue ethics. *Nous* 32, 504–530
8 Bargh, J.A. *et al.* (1996) Automaticity of social behavior: direct effect of trait construct and stereotype activation on action. *J. Pers. Soc. Psychol.* 71, 230–244
9 Keizer, K. *et al.* (2008) The spreading of disorder. *Science* 322, 1681–1685
10 Nisbett, R. and Wilson, T. (1977) Telling more than we can know: verbal reports on mental processes. *Psychol. Rev.* 8, 231–259
11 Isen, A.M. and Levin, P.F. (1972) Effect of feeling good on helping: cookies and kindness. *J. Pers. Soc. Psychol.* 21, 384–388
12 Doris, J.M. and Murphy, D. (2007) From My Lai to Abu Ghraib: the moral psychology of atrocity. *Midwest Stud. Philos.* 31, 25–55
13 Darley, J.M. and Batson, C.D. (1973) From Jerusalem to Jericho: A Study of Situational and Dispositional Variables In Helping Behavior. *J. Pers. Soc. Psychol.* 27, 100–108
14 Doris, J.M. (2002) *Lack of Character: Personality and Moral Behavior.* Cambridge University Press
15 Wilson, T.D. (2002) *Strangers to Ourselves: Discovering the Adaptive Unconscious.* Harvard University Press
16 Harman, G. (1999) Moral philosophy meets social psychology: virtue ethics and the fundamental attribution error. *Proc. Aris. Soc.* 99, 315–331
17 Bargh, J.A. (2008) Free will is un-natural. In *Are We free?: The Psychology of Free Will* (Baer, J. *et al.*, eds), pp. 128–154, Oxford University Press
18 Wegner, D.M. (2002) *The Illusion of Conscious Will.* MIT Press
19 Appiah, K.A. (2008) Experimental philosophy. *Proc. Am. Phil. Ass.* 82, 7–22
20 Buchanan, M. (2009) Secret signals. *Nature* 457, 528–530
21 Baumeister, R.F. (2005) *The Cultural Animal: Human Nature, Meaning, and Social Life.* Oxford University Press
22 Dalley, J.W. *et al.* (2004) Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. *Neurosci. Biobehav. Rev.* 28, 771–784
23 Eagle, D.M. *et al.* (2008) Stop-signal reaction-time task performance: role of prefrontal cortex and subthalamic nucleus. *Cereb. Cortex* 18, 178–188
24 Tomasello, M. *et al.* (2003) Chimpanzees understand psychological states – the question is which ones and to what extent. *Trends Cogn. Sci.* 7, 153–156
25 Crockford, C. *et al.* (2007) Baboons eavesdrop to deduce mating opportunities. *Anim. Behav.* 73, 885–890
26 Sapolsky, R.M. (2002) *A Primate's Memoir.* Simon & Schuster
27 Bembenutty, H. and Karabenick, S.A. (2004) Inherent Association Between Academic Delay of Gratification, Future Time Perspective, and Self-Regulated Learning. *Educ. Psychol. Rev.* 16, 35–57
28 Mischel, W. *et al.* (1989) Delay of gratification in children. *Science* 244, 933–938
29 Blair, C. and Diamond, A. (2008) Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Dev. Psychopathol.* 20, 899–911
30 Diamond, A. (2006) The early development of executive functions. In *Lifespan, Cognition: Mechanisms of Change* (Bialystok, E. and Craik, F., eds), pp. 70–95, Oxford University Press
31 Förster, J. *et al.* (2005) Accessibility from active and fulfilled goals. *J. Exp. Soc. Psychol.* 41, 220–239
32 Zeigarnik, B. (1927) Das behalten erledigter und underledigter handlungen. *Psychol. Forsch.* 9, 1–85
33 Aron, A.R. *et al.* (2007) Converging evidence for a Fronto-Basal-Ganglia network for inhibitory control of action and cognition. *J. Neurosci.* 27, 11860–11864
34 Miller, E.K. and Cohen, J.D. (2001) An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202
35 Rushworth, M.F.S. *et al.* (2004) Action sets and decisions in the medial frontal cortex. *Trends Cogn. Sci.* 8, 410–417
36 Fuster, J.M. (2008) *The Prefrontal Cortex.* (4th edn), Academic Press
37 Damasio, A.R. (1994) *Descartes' Error.* Putnam
38 Damasio, A.R. *et al.* (1991) Somatic markers and the guidance of behavior. In *Frontal Lobe Function and Dysfunction* (Levin, H. *et al.*, eds), pp. 217–229, Oxford University Press
39 Bechara, A. *et al.* (1994) Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 50, 7–12
40 Bechara, A. *et al.* (1996) Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cereb. Cortex* 6, 215–225
41 Koob, G.F. (2006) The neurobiology of addiction: a neuroadaptational view relevant for diagnosis. *Addiction* 101 (Suppl. 1), 23–30
42 Koob, G.F. and Le Moal, M. (2008) Addiction and the Brain Antireward System. *Annu. Rev. Psychol.* 59, 29–53
43 Frankle, W.G. *et al.* (2005) Brain serotonin transporter distribution in subjects with impulsive aggressivity: a positron emission study with [^{11}C]McN 5652. *Am. J. Psychiatry* 162, 915–923
44 Beitchman, J.H. *et al.* (2006) Serotonin transporter polymorphisms and persistent, pervasive childhood aggression. *Am. J. Psychiatry* 163, 1103–1105
45 Nelson, R.J. and Trainor, B.C. (2007) Neural mechanisms of aggression. *Nat. Rev. Neurosci.* 8, 536–546
46 Ferrari, P.F. *et al.* (2005) Serotonin and aggressive behavior in rodents and nonhuman primates: Predispositions and plasticity. *Eur. J. Pharmacol.* 526, 259–273
47 Sorrells, S.F. and Sapolsky, R.M. (2007) An inflammatory review of glucocorticoid actions in the CNS. *Brain Behav. Immun.* 21, 259–272
48 Lupien, S.J. *et al.* (2007) The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain Cogn.* 65, 209–237
49 Aarts, H. and Dijksterhuis, A. (2000) Habits as knowledge structures: automaticity in goal-directed behavior. *J. Pers. Soc. Psychol.* 78, 53–63
50 Wu, T. *et al.* (2004) How self-initiated memorized movements become automatic: a functional MRI study. *J. Neurophysiol.* 91, 1690–1698
51 Wu, T. (2008) Modifications of the interactions in the motor networks when a movement becomes automatic. *J. Physiol.* 586, 4295–4304
52 Poldrack, R.A. *et al.* (2005) The neural correlates of motor skill automaticity. *J. Neurosci.* 25, 5356–5364
53 Fincham, J.M. and Anderson, J.R. (2006) Distinct roles of the anterior cingulate and prefrontal cortex in the acquisition and performance of a cognitive skill. *Proc. Natl. Acad. Sci. U. S. A.* 103, 12941–12946
54 Reichle, E.D. *et al.* (2000) The neural bases of strategy and skill in sentence-picture verification. *Cognit. Psychol.* 40, 261–295
55 Landau, S.M. *et al.* (2004) A functional MRI study of the influence of practice on component processes of working memory. *Neuroimage* 22, 211–221
56 Sayala, S. *et al.* (2006) Increased neural efficiency with repeated performance of a working memory task is information-type dependent. *Cereb. Cortex* 16, 609–617
57 Baumeister, R.F. *et al.* (2007) Self-regulation and the executive function: the self as controlling agent. In *Social Psychology: Handbook of Basic Principles* (2nd edn) (Kruglanski, A. and Higgins, E., eds), pp. 516–539, Guilford
58 Baumeister, R.F. *et al.* (2007) The strength model of self-control. *Curr. Dir. Psychol. Sci.* 16, 396–403
59 Aron, A.R. *et al.* (2004) Inhibition and the right inferior frontal cortex. *Trends Cogn. Sci.* 8, 170–177
60 Aron, A.R. and Poldrack, R.A. (2006) Cortical and subcortical contributions to stop signal response inhibition: role of the subthalamic nucleus. *J. Neurosci.* 26, 2424–2433