



**UWL REPOSITORY**  
**repository.uwl.ac.uk**

Free will and epistemology: a defence of the transcendental argument for freedom

Lockie, Robert ORCID: <https://orcid.org/0000-0001-8340-5158> (2018) Free will and epistemology: a defence of the transcendental argument for freedom. Bloomsbury Academic, London and New York, pp. 67-106. ISBN 978-1-3500-2904-0; 978-1-3500-2906-4; 978-1-3500-2905-7

This is the Accepted Version of the final output.

UWL repository link: <https://repository.uwl.ac.uk/id/eprint/4195/>

**Alternative formats:** If you require this document in an alternative format, please contact: [open.research@uwl.ac.uk](mailto:open.research@uwl.ac.uk)

**Copyright:**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy:** If you believe that this document breaches copyright, please contact us at [open.research@uwl.ac.uk](mailto:open.research@uwl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

## Epistemic Agency and Executive Control

### 4.1 Grouping executive functions

Diamond (2013: 136) boldly claims ‘there is general agreement that there are three core [executive functions]’ – (1) working memory (sometimes ‘updating’), (2) cognitive flexibility (sometimes ‘shifting’) and (3) inhibition. These may be used to group sub-types in a typical factor-analytic hierarchical fashion (say, inhibition as including under it *suppressing attention, impulse control, self-regulation*, etc.). For exegetical/heuristic purposes, I shall partly follow her scheme in the three sections that follow, whilst rather strongly disagreeing with her claim of ‘general agreement’.<sup>1</sup> I would caution the reader: *do not look at this ordering with the critical, analytic, destructive faculties of a philosopher appraising a taxonomic scheme*. Considered thus, this ordering, and any alternative currently available, is shot full of holes: category errors; putatively distinct and fundamental executive functions (EFs) needing to appeal to some other, logically prior EF in order to function; higher levels needing to appeal to lower; one concept segueing into another; agential-level constructs (‘planning?’) needing to be used fairly gesturally and invoking several processes acting in concert. There are at least four clusters of good reasons for this state of affairs. In addition to the most obvious – the need for greater progress at the level of constructive scientific discovery – there are also major theoretical issues needing to be settled. These include deep matters regarding localization versus its alternatives; faculty psychology versus its alternatives; reductionism versus its alternatives; the sociocultural permeability of cognition; homuncularism versus its systems-based and functionalist alternatives; extended mind issues and two-process theory issues. These issues are discussed, at least to a depth appropriate to our epistemic needs, in what follows. Thirdly, there are levels of explanation/generality/operationalization/scope issues. EFs show great variegation in generality and scope. They may involve the agential level or process level (and

probably several kinds of process level). They may be measured by short time-interval, cold-cognitive experimental operationalizations (the tripartite ordering Diamond refers to was based on the use of such operationalizations as inputs to factor analysis); but they may extend over very long intervals also (e.g. 'planning') and may (characteristically, *do*) involve socioaffective thought ('hot' cognition). Thus, for example, the narrow-scope, cold-cognitive operationalizations of EF segue into things like cognitive–personality variables.

Finally, and of the first importance, there is the reflexive nature of these processes more generally: that they apply to each other, often several times over. So, *executive attention* is an executive process, and *monitoring for errors* is an executive process; but one may *direct one's attention* as to how one is *monitoring* a cognitive process for errors, or one may *monitor for errors* how one is *directing one's attention*. And this kind of reflexiveness is true for all of these processes, often to an exquisite degree (which might tell us something rather attractive about the nature of some of the aforementioned Big Theoretical Issues actually – regarding, say, self-organization, holism, emergentism, the agential level, downwards causation and non-homuncular control).<sup>2</sup> These are all of them metacognitive/metapsychological capabilities and they are all of them putatively control *processes* – they are active and ongoing, and interpenetrate/alter each other in real time. Nevertheless, we have to start somewhere, so over the next three sections we will loosely retain sight of this tripartite organizing scheme, whilst in its light (and, gradually, in light of the psychological literature more generally) considering the nature and extent of our actual powers of intellectual control and cognitive agency.

## 4.2 Working memory

Working memory (WM) involves holding things in mind and manipulating them – whether these things be visuospatial, mathematical, linguistic or other. WM as a concept grew out of Miller's (1956) short-term memory (STM: though STM is now seen as far more restrictive a process; for example, one more involved in simple maintenance rehearsal than in cognitive operations upon the items rehearsed). Miller's (1956) was then followed in 1960 by Miller et al.'s seminal *Plans and the Structure of Behaviour* (concepts from this model we will employ in an example below). WM was first explicitly named thus in the hugely influential Baddeley and Hitch (1974) model – with its *central executive*, *phonological loop* and *visuospatial scratchpad*. It has become clear that WM is one of the most

important concepts in cognitive psychology of the last 40–55 years – perhaps the most important. As this literature expanded, the notion of WM became ever more elastic<sup>3</sup>: indispensable, yet something of a monster.

Theories of WM kept developing, with this notion becoming (at times) all things to all people. Miyake and Friedman (2012) see WM (they call it ‘updating’) as including under it *monitoring*, *item addition*, *active maintenance* and *item deleting*, but for the reasons articulated already, any such list should undoubtedly be seen as far more illustrative than definitive. WM, as for all the EFs, is a limited-capacity system. It is used in mentally ordering and reordering items, in holding information and working with it, in updating one’s information, in translating goals into action plans. It is involved in all forms of analysis and subsequent synthesis: in recombining elements to form new or different wholes. Any form of reasoning involves WM. Working with any items that are perceptually or temporally no longer present involves WM. Transforming and manipulating any items in a mental ‘workspace’ involves WM.

In any actual application of WM, it will operate in conjunction with other EFs, and our examples in what follows will therefore not much flag this when it occurs. Presently, two models of WM predominate: Cowan’s, and successors to Baddeley and Hitch’s – with a good review to be found in Conway et al. (2011); but for our philosophical purposes, we can start with an example using a much earlier model.

Imagine yourself an agent instantiating a simple test–operate–test–exit (TOTE) strategy in WM – of the kind articulated in Miller et al.’s classic (1960) – itself highly mediated by the dorsolateral prefrontal cortex (DLPFC). (You may substitute a more recent theory should you so desire.) You are an inexperienced, recently qualified driver about to pull away from the kerb – at about that period in a young driver’s career when actions are starting to become highly proceduralized, and good, or bad, habits are becoming ingrained. Your safety drills are not yet (fully) proceduralized (so, this sequence is occurring in WM, not just lower brain structures, and is occurring in partly conscious WM at that). You bring up from long-term memory the mantra you learnt rigorously in order to pass your driving test: *mirror, signal, manoeuvre* – you have not yet gone the way of so many unconscientious young drivers and allowed this to pass into desuetude.<sup>4</sup> You keep in mind (‘maintain’) the superordinate goal (pulling away from the kerb) and you sequence the sub-goals. You TEST (not mirror-checked yet), you OPERATE (mirror check), you TEST (not signalled yet) ... and so on, potentially all the way to EXIT (the motor operation: pulling away from the kerb). But you are attention-sharing with a vague and nagging preconscious

awareness of the need to insert a *check for blind spots* somewhere in the *mirror, signal, manoeuvre* drill (explicit reasoning and a conscious effort at retrieval from declarative memory would, if actioned, place it between *signal* and *manoeuvre*). Implementing the drill even to this point involves (for an inexperienced driver) relatively high-order, effortful cognition in WM. *Choice between strategies* is, however, a very high-order cognitive task, and it incurs specific cognitive costs. It is widely but not unanimously agreed (cf. Miller and Cohen 2001) that such metacognitive decisions are mediated by more anterior locations in the prefrontal cortex (PFC) than the first-order EFs (here: TOTE activity).

Like Koechlin [Koechlin, Ody and Kouneiher (2003) – an alternative, also hierarchical model], they [Badre and D’Esposito (2009)] recognize that posterior regions [in the DLPFC] support the formation and execution of simple rules linking stimuli to behavior ... But they contend that more anterior regions support higher-order policies for behavior, which are needed to determine which of several simple rules applies in the current context. (Purves et al. 2013: 452)

We see an ascent in prefrontal activation from the lower (but still highly sophisticated, still characteristically prefrontal) kinds of cognition in the posterior regions of the PFC to the highest, most uniquely human metacognitive abilities in the most anterior. And this hierarchical ascent is found in both the DLPFC and the neighbouring dorso-medial PFC (DMPFC)<sup>5</sup>:

... increasingly complex sorts of control processes evoke activation in progressively more anterior regions within DMPFC ... specifically ... response-related control [control under the influence of stimulus contingencies] evokes activation in the most posterior region, the difficulty of choosing between two options (i.e. decision-related control) evokes activation in a middle region, and deviations from an ongoing decision strategy in order to adopt a new decision strategy (i.e. strategy-related control) evoke activation in the most anterior region. (Purves et al. 2013: 457)

This choice, between strategies, is partly optional for our driver (indeed, the choice whether to continue using the simpler rule *mirror, signal, manoeuvre* was optional after the driving test had been passed – it is a classic ‘tracing case’ (Kane 1996; Aristotle 2000: III, v, 1114a) that many a young driver falls into irresponsibly lax but then hard-to-shift habits in this way). We are talking about decision-related and strategy-related *control*. The agent may or may not regularly implement either the simpler rule, or no rule, or search for (and implement) more complex rules – like *mirror, signal, check blind spot, manoeuvre*.

An awareness of the need to select, then apply, the more complex rule is, as noted, in this example stipulated to be pre-conscious (hovering at the edges of consciousness, sharing finite WM and attentional resources in a classic P-FIT<sup>6</sup> manner). Finding then implementing this further rule involves the application of reasoning procedures and context-dependent information not simply in the form of an over-learned mantra or production system of rules ('*mirror, signal, manoeuvre* but include *check for blind spots* when pulling away from the kerb ... or overtaking on the motorway ... or ...'). Serious attentional demands occur when having to (a) perform the motor task itself, (b) perform a TOTE check in executive monitoring of the task and (c) effect high-level choice of strategy, and strategy implementation, at the same time.

You are just at the stage of driving experience where you could enjoy cutting free from some of those attentional costs through automatization – and you *can* perform this pulling away at lower attentional cost and consequently greater motor fluency. Perhaps even there is someone in the car whom you wish to impress with said greater motor fluency. But you perform the slow, awkward, attentionally burdensome sequence of checks, and checks for context-relative strategy upgrades, and consequent higher-order strategy selection, then implementation, and further checks, then initiate this strategy through sequences of further TOTE testing – perhaps it takes an additional second or two, perhaps the motor sequence is in consequence slightly more jagged. You get to the 'check blind spot' over-the-shoulder motor procedure just prior to EXIT (to motor driving commands) and manage a very ballistic saccade just as a cyclist whizzes past your driver's side wing. You are lucky, and he is luckier still. But he is only lucky. You are diligent also. You achieved knowledge (cyclist, beware!) through dutiful, effortful cognition (effortful control, effortful attention, themselves examples of conscientiousness – and related high-level cognitive–personality variables much studied by individual differences psychologists). You are deontically epistemically justified. Over time, this will have a 'tracing effect' on your character as a driver – and these procedures will become automatized into lower-brain routines. Clifford, that delicious *enfant terrible*, would have immediately appreciated this example as epistemology, though he might have struggled with the neuropsychology (and perhaps have been more conversant with a ship than a car).

What, then, is meant by the claim that this belief ('there is a cyclist there, I must not pull out!') is *involuntary* in some sense that warrants the conclusion: *hence not deontically epistemically appraisable*? What, further, is meant by the claim or assumption that its involuntary status (in this sense) is the clear

verdict of modern psychological science, or is secure sans any such science, merely from the armchair? The afferent (final) visual process is in this case involuntary, granted – it is a triggered, mandatory output of the visual system. Moreover, in this case, that we perform the final efferent saccade is triggered and involuntary also – it obeys Sechenov’s principle: thought is not the cause of this action, the stimulus (cyclist) is. Such reflex eye movements (‘express’ rather than ‘intentional’ saccades) are mediated more exclusively by the superior colliculus than the higher brain centres in the frontal eye fields (of which more later). They are examples of *involuntary* (rather than *executive*) attention. But can we really rest with this and say the saccade is efferent and involuntary (and perhaps thereby not really actional or at least not agential), and the visual processing consequent upon that is afferent, triggered, involuntary; while the actional look over the shoulder is efferent and voluntary but exclusively motor-actional, and subject thereby to merely ethically rather than epistemically deontic appraisal? Can we say there is no choice or control at the epistemic level because the final processes of afferent visual cognition and efferent cognitive–motor attention are triggered, mandatory, modular? Can we say that *cognitively* no actions are taking place here? Such a contention would seem to me entirely indefensible. The separating out of the final pathway as alone epistemic, of some unique importance, is nowhere defended in cases like this and would appear entirely wanting for motivation. So, some final pathway of low-brain afferent sensory processing is involuntary – so what? We are normatively assessing all the important, distinctively human cognition that occurred prior to that; or (perhaps better) we are deontically assessing that lower-brain automatic processing only inasmuch as it is integrated with, and tasked by, our higher, controlled cognition: only inasmuch as it is owned by the agent, who is responsible for it. A point of great emphasis in the deontological epistemic tradition is that the process of cognition would have been commendable, and its absence remiss (as for Clifford’s judgement on the lucky but culpable owner of an unseaworthy ship), even had the cyclist not been whizzing past the driver’s side wing, thus no final, proximal process of afferent lower-brain cognition *at all*:

Let us alter the case a little, and suppose that the ship was not unsound after all; that she made her voyage safely, and many others after it. Will that diminish the guilt of her owner? Not one jot. When an action is once done, it is right or wrong for ever; no accidental failure of its good or evil fruits can possibly alter that. The man would not have been innocent, he would only have been not found out. (Clifford 1999: 71)

It is *not* the final, fast, involuntary, afferent sensory processing that we are normatively epistemically assessing here; it is not and never was thus, and the opponents of the deontological tradition can hardly have been unaware of this. The perception of the cyclist involved you responsibly *directing* your thought. Your conduct in this case is deontically epistemically justified because you directed first your inner thought, and then your outer motor–cognitive actions – the look over the shoulder – diligently, thus to be epistemically justified. You *controlled your cognition*, and did so well: you discharged your intellectual duty, thereby to acquire knowledge.

### ***Two-stage and final pathway arguments (again)***

Our driver–cyclist case was a simple example of WM involving a series of mental acts eventuating in a motor efferent and consequent sensory/perceptual afferent; one that left space for the (I maintain) illicit, philosophically contrived separation into two stages<sup>7</sup>: the EF stage (typically seen by involuntarist epistemologists in terms of a gestural, armchair model of attention or something like, and typically, more-or-less grudgingly conceded to be controlled cognition of a sort) and the triggered ‘final pathway’ sensory/perceptual stage, seen (correctly in most cases, at least of this simple a nature, and at least if allowed to be *final enough*) as involuntary. At the point in the last chapter when this two-stage separation was introduced, it should already have been apparent that affecting to see the second ‘final pathway’ stage as voluntary cognition is something strikingly few actual deontologists were or are tempted to do; and further, that the first stage is all one needs and all that most actual deontologists have ever wanted as the basis for deontological appraisal. Noting the appropriateness of deontic evaluation of the ‘first-stage’ regulative cognition in the absence of the final pathway (when there is no cyclist – Clifford cases) is one way to make this point. Note, however, we also and especially need normative epistemology for cases of very important practical or pure theoretical reasoning where there is in no obvious sense any ‘final process’ of mandatory, triggered, sensory processing at all.

### ***Cognition without any clear ‘second stage’: Heisenberg cases***

WM is involved at every level of higher cognition, including (especially) cognition with no significant sensory afferents and no significant motor efferents – cognition intrinsically rather than accidentally lacking this low-level, typically sensory ‘final pathway’ – cognition that includes the highest cognitive



achievements of mankind. Consider cognition unprompted by any efferent involuntary saccade or afferent ‘triggered’ sensory processing or any proximal efferent motor output, voluntary or otherwise: endogenously guided, stimulus-independent thought. Consider ratiocination about a practical matter of great personal concern,<sup>8</sup> or deep thought about a theoretical matter as an end in itself. The same attentional and/or executive costs have to be paid – and then some: the resolve needed ‘to drill in very hard wood ... and go on thinking beyond the point at which thinking begins to hurt’ (Heisenberg, in Powers 1993: XV).

Let us take as our agent, then, Heisenberg, thinking, purely for its own sake, and beyond the point at which thought has started to hurt. Granted, he’s working with a pencil and notebook, and moving his eyes (of which more when we reach matters pertaining to the situated/embodied cognition literature), but he’s surely doing a lot of thinking not involving any motor–sensory processing at all. Suppose there is an over-learned, obvious response to the elements of the mathematical task in front of him (a ‘mental set’ – a Gestalt *Einstellung*). It would be awfully hard maths for even a postdoctoral specialist, but by now it’s relatively straightforward for Heisenberg, and by now he suspects it’s the wrong maths. He has made many attempts to think his way past this impediment and avoid the prepotent response (simply to go down the over-learned, for-him cognitively straightforward, ‘functionally fixed’ response route). So Heisenberg ruthlessly maintains the elements of his mathematical cognition in WM at considerable cognitive cost, refusing to be led by easy cognitive-cue ‘lures’ into distraction – high levels of ‘goal shielding’ are consciously and deliberately applied.<sup>9</sup> In so doing, Heisenberg forcibly prevents a prepotent response set from being implemented (he prevents ‘entrenchment’). Such cases involve ‘active maintenance and management of task goals in the face of interference’ (Friedman et al. 2008: 219). Heisenberg facilitates inhibitory control (of the *Einstellung*) by effortfully keeping the cognitive elements in WM. Thus does he escape ‘functional fixedness’ and discover, in a moment of Jamesian–Gestaltist insight,<sup>10</sup> an innovative *reductio* solution to his problem. It hurts, but it’s commendable, and it’s WM in the service of inhibitory control.

Notice, however, that in this example (of WM in the service of inhibitory control), Heisenberg also, and at the same time, needed to do the reverse, and use inhibitory control (effortful, specifically inhibitory attention) to achieve this mastery over WM (the reflexivity/promiscuity/task impurity motif that pervades EF). He needed to override prepotent responses, resisting proactive interference in order to forcibly retain material in WM when he’d rather chunk it, or sequence forward from it into the more obvious maths – or glaze his eyes over

into the middle distance. And his awareness (hard earned) that he had to do this was metacognitive awareness, born of repeated failure at this task, coupled with Bartlett–Neisser top-down schematic expertise from experience with many such tasks (itself deontically and aretaically commendable, the product of decades of effortful cognition). We have *metacognition*, *inhibitory control*, *executive attention*, *WM*, *flexibility*, *crystallized IQ* (his mathematical expertise, his schemata), and *sheer bloody effort* (both at this time and in the past) all combining elegantly, systemically, in very high-level, regulative, agential cognition. That is executive functioning: it is both the most controlled and the highest-order cognition we have; it gave us the greatest cognitive accomplishments of mankind; and its levels of control both permit and encourage deontic epistemic appraisal.

### 4.3 Cognitive flexibility

‘Cognitive flexibility’ involves such notions as task shifting, set shifting, ‘creativity’,<sup>11</sup> divergent thought, taking the character of the other, theory of mind, self-awareness, strategy selection, mental fluency, linguistic fluency/internalization of speech (as noted, sometimes separated out as a distinct factor). To reiterate: all EFs are incipiently metacognitive – but certainly this family of EFs is. It, too, may be tested with low-level, cold-cognitive operationalizations (e.g. card-sort, trail-making, alternative uses); yet once again these exist in interaction with certain high-level, cognitive–personality variables. Cromwell’s famous ‘I beseech you, in the bowels of Christ, think it possible that you may be mistaken’ indicates (I take it) both the *responsibility* and the *flexibility* aspects of the case I am seeking to make in normative epistemology. Cromwell presumably felt his invocation *could be acted upon* by the members of the synod of the Church of Scotland: that they had the metacognitive power to entertain the possibility that they might be wrong – to exercise their responsible judgement and shift the parameters within which their first-order cognition was taking place upon being addressed in this fashion, perhaps in a classic ‘gain’/bias fashion, perhaps through set shifting, perhaps though ‘taking the character of the other’/theory of mind capabilities, perhaps in various of the ways the classic problem-solving literature identifies (a change in schema/Gestalt), perhaps in some other way. We hold people responsible for dogmatism, small-mindedness, auto-indoctrination, rigidity and inflexibility of thought. We commend people for ‘walking a mile in the other man’s moccasins’, for creativity and open-mindedness and divergent thought. We may operationalize these things as cognitive–personality variables quite

apart from the more narrowly cold-cognitive operationalizations of flexibility made famous in the experimental-cognitive EF literature. Metapsychologically, to return to our reflexivity motif, individual differences in *conscientiousness* relate specifically to individual differences in open-mindedness/flexibility EFs (Fleming et al. 2016 – though with direction of cause not established) a relation, I take it, that Cromwell's plea trades upon.

In the previous example, Heisenberg shifted sets in refusing to be functionally fixed on the existing family of mathematical techniques he had used to that point. Famous Gestalt examples from the problem-solving literature could of course be invoked here (Maier's 1930, 1931 two-string problem,<sup>12</sup> Duncker's 1945 candle-matchbox problem, etc.). As this literature copiously indicates, we have the capacity to restructure the elements of a problem, to employ analogical transfer and other techniques in order to avoid functional fixedness and to escape set effects. We also have capacities for insight that defy current mechanistic capture. The converse of task/set shifting is *staying on task* (an inhibitory EF). We saw already that maintenance of mental/attentional set (e.g. in the face of distraction, attentional lures, etc.) is an essential EF – continuing to attend to the task in hand, to 'gate off' the contents of WM and not be distracted. Frontal lobe damage profoundly impairs this ability. But frontal lobe damage also profoundly impairs the ability to shift task – with the most well-known example being the extraordinary clinical symptom cluster of *perseveration*. Lesser cases of cognitive inflexibility are measured by clinical or experimental tests like the famous Wisconsin Card Sort Test: the patient can swiftly learn to sort the cards by one rule (shape or colour or number), but cannot shift to sorting by another. Sometimes, in a very strange phenomenon, they can state that the rule has changed yet cannot change their sorting method consequent upon this.

Notice, however, that these two families of EF (task shifting, an example of flexibility; and staying on task, an example of inhibition), if described mechanistically, are simply and directly at odds with each other.<sup>13</sup> What, then, as undamaged agents, should we do in response to the demands of some given cognitive task? A deadpan refusal to acknowledge the problem here would be to answer thus: stay on task *if that is appropriate*, and task shift *if that is appropriate*. But when is the one EF 'appropriate' and when the other? Well, EFs exist precisely to permit us to flexibly address novel, context-variant problems without fixed parameters (other tasks may be solved far more efficiently with lower-brain structures' fast, evolutionarily old, inflexible algorithms). And EFs are reflexive: they apply to the application of EFs, as well as to the first-order

problem that is set. To decide whether you should stay on task or task shift, you must *use your judgement* (a gestural, non-process-level composite for EFs, including these two EFs, acting in concert at the agential level in response to the demands of the problem set). But the danger of regress is obvious: which meta-level mechanism decides when to deploy the one EF and when the other? Deep issues emerge. One could supply a straight answer here – and find a mechanistic proximal determinant (and in this particular case, narrowly conceived, perhaps we would find some token algorithm<sup>14</sup>). Actually, as we noted in considering WM, some (but not all) current models (hierarchical models like Badre and D’Esposito’s and Koechlin’s) show an ascent in prefrontal activation from the lower (but still highly sophisticated, still characteristically prefrontal) kinds of cognition in the posterior regions of the PFC to the highest, most uniquely human metacognitive abilities in the most anterior. But if you keep adding actual metacognitive process mechanisms in an ever-more anterior location, there comes a time when you burst out of the front of the forehead.<sup>15</sup> And anyway, EFs are precisely needed when mechanistic (automatized or old brain) approaches are inadequate and a creative, novel, insightful approach is needed. So what else could we say here? One could claim ‘justification comes to an end’. One could claim we have come to the limits of mechanism. Or one could start by noting that these EFs are all of them (even the more anatomically posterior and behavioural) nevertheless incipiently metapsychological, reflexive, systemic. Miller and Cohen (2001) in their very influential review, advance the theory that the PFC doesn’t so much engage in self-contained ‘executive processing’ in the sense of being a cognitive supervisor, a discrete homuncular controller-in-a-box. Rather, they see EF as *directing* and *biasing* other cognitive processes throughout the brain (cf. Section 4.8 below). Miller and Cohen use the old metaphor of a train-track controller, moving points switches to prevent collisions and to make best use of the rail network (see also a deep paper by Burgess et al. 2006). But how may we have the functions of *control* without a *controller*? Fuster (2013, 2004), after Sperry and Gibson, invokes here the notion of repeated applications of multi-level, massively interpenetrating *perception–action cycles*, with control emergent out of these cycles operating via self-organizing, broadly cybernetic open-system feedback principles.

The perception–action cycle is the circular flow of information from the environment to sensory structures, to motor structures,<sup>16</sup> back again to the environment, to sensory structures, and so on, during the processing of goal-directed behavior. (Fuster 2004: 143)

Such a picture of things, though moderately well anchored in present science, will undoubtedly require far more progress before it, or some successor, may be considered any kind of consensus view. Whatever may be the truth about this conception of the issues (and to me, it, or something like it, looks to be the most currently promising conception of matters), descriptively, I'd suggest we agree that the PFC doesn't necessarily contain some, most anterior, homuncular, *causa sui* controller. But it need not thereby be merely a C. D. Broad-style mechanistic *component* system (say, massive modularity, in the modern parlance). Rather, by any of several means, the PFC (along with other higher-brain structures) may form a distributed, marvellously coordinated, reflexively applicable *system of control*: directing and reacting to (and continuous with) other, lower-brain resources<sup>17</sup> – and, we shall see, not only *brain* resources. Regulative intellectual control is a product of the systemic whole. Mechanism does to some extent apply to the parts; but (to be argued more completely in Chapter 9) agency is emergent out of the whole.

#### 4.4 Inhibitory control

'Inhibition' is an umbrella term for many kinds of downwards control processes – some very high level. Some simpler cases of effortful inhibitory attention we have already considered regarding WM and flexibility. At the level of common and surprisingly useful operationalization, an example of inhibitory attention might be when one attends to the ink colour and actively inhibits naming of the colour word in a Stroop task ('BLUE' printed in green ink). This segues into inhibitory self-control more generally – as when one stays 'on task' despite temptations to allow one's attention to wander or 'task switch'. This inhibition is *directed*, and it is also *hard* – there are *costs* involved (one *pays* attention – the psychologists operationalize these costs in participants' performance). It is *felt* as hard. In light of the aims of this chapter, exercising such inhibitory control in a particular task may be *commendable* and failing to do so may be *remiss*. Pick a case of inhibitory attentional control; for example, a countersaccade task (looking left when the stimulus indicates right because you are required to ignore the right-hand lure). This is a classic operationalization of EF: it is a clinical and experimental test of the strength of your voluntary mental powers – to override and inhibit your involuntary attention. If you are a frontal lobe patient you cannot do this (thus cannot be blamed for not doing this) and if you aren't you can, but it will involve cognitive costs: objective (in terms of reaction time and

task impairment) and experiential (it *feels* hard, you'd rather avoid it). Pick a case where one discharges one's obligation to effortful attentional control in a matter of practical or theoretical importance (bridge lookout duties; or difficult reading; or diligent and timely discharge of an unwanted academic refereeing task – something involving directed attention, vigilance and effortful inhibitory control in order to 'stay on task' and avoid distractions). These are classic cases of *commendable* mental/motor actions, are they not? Failure to discharge such effortful cognition may then be a classic case of *remiss* mental action (/omission).

***Inhibitory control segues into a large tranche of broader inhibitory variables, cognitive–personality variables and 'hot' cognition***

Inhibitory control is not only, or even chiefly, a prescribed, 'cold-cognition' function or family of functions. Although narrowly focussed, short-duration cognitive tasks may be used to operationalize inhibitory processes, both clinically and experimentally, these are part of a large – perhaps the largest – family of EFs; and this family of inhibitory functions segues into very important cognitive–personality variables and into values-based socioaffective cognition. We possess very important effortful control capabilities. 'Effortful control' is, strictly, just one named version of a cluster of terms/factors/operationalizations that, though invoked by slightly different literatures, partially overlap and for the most part measure close to, though not exactly the same thing. Some of these involve sheer effort – willpower (which concept Roy Baumeister has done so much to rehabilitate) or, more narrowly, effortful attention, as mediated, say, by the posterior parietal lobes. The phenomenon of effortful attention by itself is surely sufficient to refute any global attempt to undermine our claims to cognitive control; but there are a wealth of such high-order inhibitory capabilities, possessed of varying degrees of generality, going well beyond the unmediated application of effort, willpower, impulse control or inhibitory attention. These may be terminologized/operationalized in various ways: as *effortful control* per se (for a review, see the excellent MacDonald 2008), as *response inhibition*, as *self-regulation* (Bjork et al. 2013), as *gratification delay* (Mischel *passim*), as *motivation*, as internality of *locus of control* (Rotter *passim*), as *consideration of future consequences* (Strathman et al. 1994), as *time discounting/impulse control*, as the *behavioural inhibition system* (Gray *passim*) – and other ways besides. For all the variability in scope, timespan, value-ladenness, level of explanation and generality that these represent, much of this literature is highly intuitive – especially to those, such as the readers of this chapter, involved in any capacity

in education. For example, the (educational) self-regulation literature articulates ways in which one may achieve some measure of self-mastery and control over one's educational development, often via metacognitive awareness and control of the learning environment. Consider a dedicated student, struggling against all the odds, mastering her discipline over a classic time period for transition to expertise (ten years or 10,000 hours: Simon and Chase 1973; Ericsson et al. 1993). Consider a great figure from our discipline, a colossus of the past, spurned by his contemporaries yet cherished by posterity. Consider just a diligently written undergraduate essay as opposed to a perfunctory 'cuttings job'. Why is deontic appraisal appropriate here? Because of the exercise of regulative control in the face of short-term cognitive costs and longer-term adversity. Because of the effort expended in the service of the agent's epistemic values: instrumental or, in many cases, ultimate. Because of the intellectual conscientiousness and integrity that has been demonstrated.

We can, to a greater or lesser extent, defer immediate gratification in the cause of higher cognitive goals. Some of this inhibitory control is mediate and some immediate. We can suppress more basic (e.g. appetitive) responses to keep our superordinate goals in mind. Our *values* most certainly inform this activity. For example, even in the rather lumpen IQ literature, many, from Binet onwards, have acknowledged the importance of what Weschler called the 'non-intellective factors' (aretaic factors: personality, motivation, character, discipline) – even for Spearman's putatively purely *intellective* factor (par excellence) of *g* (general intelligence) and even for Cattell's even purer distillation of that, *fluid intelligence*: *g*(F). There is much acknowledged evidence that dependency relationships/highly diachronic feedback loops here go in several directions at once:

... individuals who are better able to suppress their emotions have higher scores than their more impulsive peers on *g*(F). ... There are at least three possible explanations for the relationship between self-regulation and higher scores on tests of *g*(F): (a) The ability to self-regulate could be a manifestation of intelligence; (b) these constructs could share common variance such that they are both affected by a third variable; or (c) self regulation could be one of the processes that facilitate the development of intelligence. There is evidence that self-control, or at any rate some set of nonintellective motivational factors, contributes not only to life outcomes but to IQ scores themselves. (Nisbett et al. 2012: 152)

One persuaded by a generally hard determinism could insist that these higher-order control factors (valuational, motivational, character-mediated) are themselves entirely determined. As will be seen in Part 2 of this book,

I am, on other grounds, not in any way persuaded by that position; but since we are trying to see if there is anything specially untenable about epistemic deontology, having to repair to a stance whereby one's doxastic involuntarism could be defended only *sub specie* a universal hard determinism would be, I take it, dialectically beside the point – indeed, as regards the specifics of the epistemic debates, an admission of defeat. One could try for some view of these higher-order control factors whereby (at some scientific level short of a sweeping metaphysical hard determinism) they are themselves the mechanistic output of some manifestly deflationary, say, additive algorithm (whereby these factors' control capacities are somehow fictive, at least *qua* genuine higher-order control). Whether as regards the affective, the conative or the cognitive, this is not the verdict of the science, at least as it stands as of now. MacDonald considers the relative inapplicability of any such deflationary, associative model as regards even the psychology of other mammals, let alone human psychology, using an existing example in the literature of an organism subject to:

... the conflict resulting from walking across hot sand to get water. For many animals, such conflict is resolved simply by the summed strength of the competing implicitly processed action tendencies (thirst versus pain avoidance) – a standard ethological account. ... For mammals, the PFC or its analogues underlie executive control of behavior that takes into account not only subcortically generated affective cues that are routed through the orbitofrontal cortex (OFC), but also sensory input and other information (e.g., learned contingencies) available to working memory. ... For humans, the information available for executive control also includes explicit appraisals of the context: Is the water potable and, if not, could it be made potable using available technology? Will taking the water provoke an attack from those who currently possess it and, if so, could this attack be suppressed? Is the water part of a sacred ritual and therefore off limits for drinking? As this example indicates, conflict occurs not only because of conflicting signals from modules; there may also be conflicts between the output of modules and symbolic representations of the context. (MacDonald 2008: 1014)

To suppose that these top-down, characteristically human, rational capabilities, capabilities for hugely sophisticated control, capacities that are a product of the staggering complexity of the PFC (in many-ways, open-system interaction with other brain structures and with social, cultural and other contexts) – capabilities that vitiate any attempt to apply an unmodified version of the summative ethological model to humans – must nevertheless themselves be merely summative in turn, would be to embrace, from the armchair, nothing



more than a mechanistic dogma. If scientific evidence were to be forthcoming for a harshly reductive algorithm, one offering some sweeping deflation of our claims to agential, values-laden intellectual control, then this would deserve our consideration; if, sans evidence, a set of mere philosophical prejudices is evinced, it does not. Deflationary epistemologists, including those making surprisingly authoritative claims as to how the sciences of cognition would have matters be, have hitherto shown no awareness that the many scientific literatures adverted to in the foregoing even exist; they might wish to assimilate the lessons of said literatures before making further confident assertions as to the non-existence or impossibility of agential, values-laden, self-regulated, higher-order, rational cognitive control.

#### 4.5 Mediated cognitive agency

Certain motifs have emerged now concerning the EFs. They are ‘promiscuous’ (they segue into and tumble over one another – ‘task impurity’ is rampant); they are reflexive; they are incipiently metapsychological (not just metacognitive); they act on lower-brain processes and more posterior cortical processes, as well as on each other; they have many levels of description, generality and temporal span (including very high-level ‘agential’ descriptions); they involve considerable ‘cultural scaffolding’; and they are values-laden. Here we transition from the last section’s discussion of inhibitory EFs to what, for want of a better word, I shall call ‘mediated’ cognitive agency and self-regulation (effected in part via manipulation of the external environment and/or self in that environment). This can bridge to take us smoothly into the next section’s discussion of situated/embodied/extended cognition and these things’ relevance to matters of cognitive control.

We considered, via our cyclist case, the distinction between executive visual attention, mediated by the frontal eyefields and under the control of the agent, and involuntary visual attention (a ‘final-process’, triggered motor efferent followed by a mandatory sensory afferent) obeying Sechenov’s principle (under the control of the stimulus) and mediated more specifically by the superior colliculus. But the EFs have a long reach, are all of them incipiently metapsychological, and act to obtain control over lower-brain cognition one way or another – including techniques of oblique, prosthetic control. Consider, then, that we may have indirect (functional) voluntary control of our avowedly *involuntary* attention. In addition to the educational self-regulation literature, there is a wealth of material on this in other self-regulation literatures, such

as in addiction research. Addicts have problems with ‘cue reactivity’ (sight of the whiskey bottle). But one may seek after and gain some prosthetic control here. I wish to stay on task, devoting myself to writing my philosophy paper and avoiding the distracting effects of online poker. I may remove the poker site’s icon from my desktop, as it is an involuntary attentional lure – leading, if not removed, to problems with ‘disengagement’, indeed, possibly to task switching (as we have seen, another, flexibility, EF, and a crucially important one, but one that is here needing inhibition). The icon is an exogenous (automatic, sensory, stimulus-driven) attentional cue, but my chosen motor action (my mouse clicks) may remove it as a lure from my visual environment. It may do this to *inhibit* an involuntary attentional lure no less than my chosen motor action – my voluntary look over the shoulder – placed an involuntary exogenous attentional cue (the cyclist) in place to *effect* my involuntary ballistic saccade in the event of a driving hazard being present.

This set of steps (removing the poker icon) is an example of *planning*, involving *goal setting* and *sub-goal setting*, subserved by decoupled thought and neo-Piagetian/theory of mind abilities – to take on the character of the other (even though the ‘other’ here is myself). Of course, I could just leave said icon in place and exercise my immediate *impulse control* – and, in the grip of a dubious metapsychological theory (‘it’s all about willpower, isn’t it!’), I might do this – but doing so will be an additional, needless, cognitive-conative burden here. ‘Pure’ (purely ‘internal’) inhibitory control is, from the first, bound up with, (partly) internalized out of, and interpenetrates, mediate, functional (e.g. behavioural) control. Thus, Mischel’s four-year-olds would in some cases physically cover their eyes to avoid seeing the marshmallow; and this was an effective (unprompted) cognitive-actional strategy, one in principle continuous with more internalized, ‘cognitive’ strategies:

... those who were most effective in sustaining delay seemed to avoid looking at the rewards deliberately, for example, covering their eyes with their hands.  
 ... Many children generated their own diversions: they talked quietly to themselves, sang, created games with their hands and feet, and even tried to go to sleep during the waiting time. Their attempts to delay gratification seemed to be facilitated by external conditions or by self-directed efforts to reduce their frustration during the delay period by selectively directing their attention and thoughts away from the rewards. (Mischel et al. 1989: 934–5)

These are examples of functional, mediated, metapsychological control. Metacognition involves ‘internal’ and proximally ‘external’ (e.g. motor actional) self-regulation, but as noted, it also involves the ability to control our intellectual

environment more generally (e.g. as for the educational self-regulation literature: our learning environment). Actions we control will partly determine what it is that is accessible to us (and the reverse of course). To employ some terms of art within Piagetian theory: processes of assimilation, accommodation and equilibration criss-cross the active–passive divide. To amplify a motif from Chapter 1: this should be seen as representing a partial rejection of *introspection* as the sole or chief criterion for accessibility in epistemology. The active and passive features of our cognition are in an inter-penetrative dialectic – a ‘perception–action cycle’ (actually: many such cycles). It would probably be a mistake to locate this access or agency at some single, Archimedean, Cartesian, introspective point. So, consider a modern Piaget-influenced theory: in Sternberg’s triarchic theory of intelligence, the Practical (Contextual) sub-component concerns the ‘external’ aspects of intelligence – the ability of the individual to (a) adapt her behaviour to her environment; but also (crucially) her ability to (b) *shape* that environment and to (c) *select* environments that suit her, or get out of environments that don’t (‘niche picking’). The (educational) self-regulation literature repeatedly emphasizes the student’s need to learn how to manage the conditions of her own learning; and the emotional self-regulation literature emphasizes ‘situation selection’ and ‘situation modification’ as the first two out of the five components of its ‘modal model’ (Gross and Thompson 2007).<sup>18</sup>

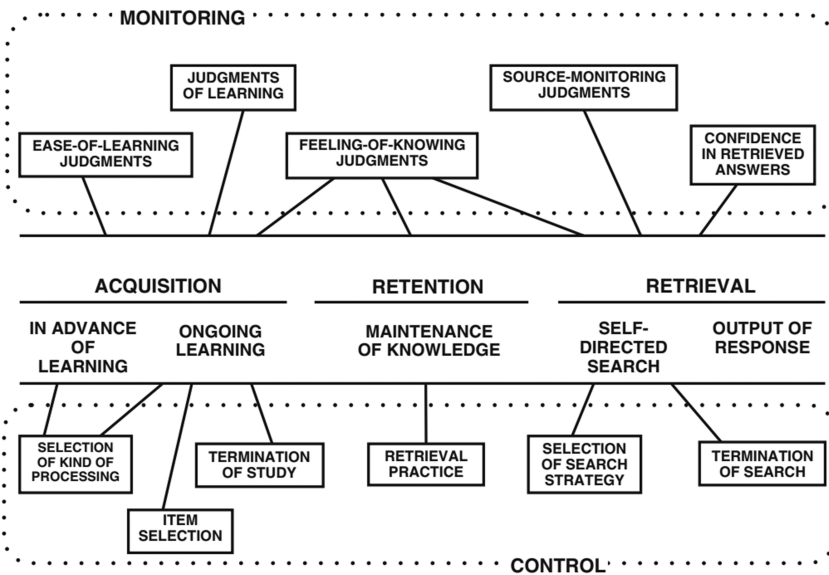
### ***Self-regulation***

The educational self-regulation (SR) literature concerns our metapsychological abilities to self-manage, control and self-initiate learning; whilst the emotional SR literature (e.g. the ‘hot sand’ example of MacDonald, considered earlier) concerns our metapsychological abilities to self-manage more generally, including socioaffectively. These are connected, as any teacher will know. Quite a lot of the educational literature pertains to memory, and here we face the difficulty that (even before considering the metacognitive level) at the first-order level, philosophers are too often in the grip of a ‘preservation’ (/*contamination*) view of memory as a *faculty* for *encoding* events/items/information – a passive and automatic ‘module’ (i.e. one separable from, albeit perhaps connected to, other cognitive capabilities), a *faculty* for *recording* information whose mandatory outputs may perhaps be ‘contaminated’ by *biases* from the rest of the cognitive system, but which otherwise *preserves the memory trace*. This picture I shall polemicize elsewhere. The whole, vast, neo-Bartlettian tradition

undermines this as a conception of memory. Here is a statement closer to the truth:

We store new information in terms of its meaning to us, as defined by its relationships and semantic associations to information that already exists in our memories. What that means, among other things, is that we have to be an active participant in the learning process – by interpreting, connecting, interrelating, and elaborating, not simply recording. Basically, information will not write itself on our memories. (Bjork et al. 2013: 420)

These claims concern first-order learning (itself agential, active, top-down – in all the ways the Bartlettian, Piagetian and Gestaltist traditions have indicated over the last 80–100 years). This first-order conception of memory is already very difficult to square with a global cognitive involuntarism. But the SR literature adds levels of metacognitive activity above this. Although the details will not concern us, for illustrative purposes, Dunlosky, Serra and Baker’s (2007) diagram for some of the complex systems of monitoring and control of encoding, retention and retrieval is given in Figure 4.1. What was (in the preceding chapter) referred to as metapsychological *knowledge* is here called ‘monitoring’, while metapsychological *regulation* is called ‘control’.



**Figure 4.1** Metamemory – monitoring and control

Source: Dunlosky et al. (2007: 141, their Figure 6.2), itself adapted from the Nelson and Narens (1990) framework for metamemory. Copyright Wiley 2007.

Acquiring these metacognitive abilities is *hard*, but these are ‘desirable difficulties’:

Becoming sophisticated as a learner requires knowing how to manage one’s own learning activities. In that respect, we are both teacher and student. What makes acquiring such sophistication difficult is that the short-term consequences of introducing manipulations such as spacing, variation, interleaving, and generating<sup>19</sup> can seem far from beneficial. Such manipulations introduce difficulties and challenges for learners and can appear to slow the rate of learning, as measured by current performance. Because they often enhance long-term retention and transfer of to-be-learned information and procedures, they have been labeled desirable difficulties (Bjork 1994), but they nonetheless can create a sense of difficulty and slow progress for the learner. (Bjork et al. 2013: 421)

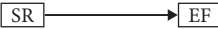
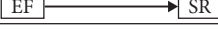
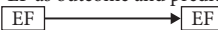
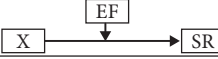

This metacognitive processing is itself controlled (EFs are reflexive). One forms what this literature calls ‘judgements of learning’ (JOLs) throughout. If these JOLs are reductive, mechanistic algorithms, then that is certainly not an assumption this literature embraces. ‘Metacognitive judgements are also inferential’ (Bjork et al. 2013: 428). The role of *monitoring* (according to Miyake and Friedman 2012, a WM or ‘updating’ EF, but according to many others and on face validity, an inhibitory/attentional EF) is found throughout successful self-regulation of learning: the more you self-monitor the better, with active retrieval (itself agential, active, self-cued) also a learning event, something that re-encodes and alters the stored memories themselves.

Our review suggests that effective learning can be fun, it can be rewarding, and it can save time, but it is seldom easy. The most effective cognitive processes involve some effort by the learner – to notice connections and linkages, to come up with examples and counterexamples, to generate and retrieve, and so forth. In short, effective learning requires the active participation of the learner. (Bjork et al. 2013: 436)

If you demonstrate this active participation, that is commendable; if you don’t, that may be through no fault of your own, but if it is your own fault, you are remiss. Self-regulation is rewarding, *but it is seldom easy*. It is, however, a characteristic of human cognition – one that invites and is highly compatible with a deontic assessment of the self-regulating cognizer.

### ***Relationship between executive functioning and self-regulation***

Hofmann et al. (2012: 176) consider various ways in which EF and SR may relate to each other. Their figure is here reproduced as Figure 4.2.

Conceptual approach	Example research questions
EF as outcome 	<ul style="list-style-type: none"> <li>• Does dealing with strong cravings consume WM resources?</li> </ul>
EF as predictor 	<ul style="list-style-type: none"> <li>• Does WMC predict mind-wandering in daily life?</li> <li>• Which EF is most strongly related to emotion regulation?</li> </ul>
EF as outcome and predictor 	<ul style="list-style-type: none"> <li>• Do separate EFs have negative aftereffects on each other?</li> </ul>
EF as process moderator 	<ul style="list-style-type: none"> <li>• Are impulsive precursors better predictors of risky sexual behavior among people low in inhibitory control?</li> </ul>
EF as process mediator 	<ul style="list-style-type: none"> <li>• Are the effects of stereotype threat on performance brought about by temporary reductions in WM capacity?</li> <li>• Does repeated training boost inhibitory control &amp; does this training gain in turn translate into improved self-control in everyday life?</li> </ul>

**Figure 4.2** Different approaches to the study of executive functioning and self-regulation

Source: Hofmann et al. (2012). Copyright Elsevier 2012.

SR may predict EF, as when our diligent student shows self-discipline, willpower and impulse control – to resist distraction, concentrate hard and thus ‘gate off’ WM from appetitive attentional lures, in order thereby to think better. EF may predict SR, as when WM capacities or inhibitory controls subserve our self-regulatory capacities. EFs may affect other EFs – the promiscuity/reflexivity motif that runs right through this chapter. Or EF may moderate/mediate a process of SR, as when Mischel’s four-year-olds generated their own attentional diversions to mediate their capacities for gratification delay in the presence of a stimulus (the ‘X’ here being the marshmallow).

Notice that this complexity is found even at the level of EFs alone (to the extent these may be distinguished from SR capacities anyway). Our capacities for hot and for cold executive, agential, self-regulative control segue into, interpenetrate, reflexively affect and feed back to one another in processes of intricate and systemic elegance.

### 4.6 Motor-actional freedom interpenetrates mental-actional freedom

Turn away from the direct confrontation with claims of involuntarism, to the contrast between a supposedly unproblematic freedom of motor-action and the specific problems supposed to attend any acceptable notion of intellectual freedom. Such a stance may be attacked via its assumption of an unproblematic categorical distinction between motor-actional and mental agency. Here we may help ourselves to a great body of work within the embodied and situated

cognition literature. Any categorical distinction between motor and mental actions simply breaks down: to cede motor-actional freedom is (at least for intact beings like us, and at least to a goodly extent) to cede mental freedom.

Consider a very familiar example from this literature: a pattern-matching task, such as might occur when playing the video game Tetris or finding a piece to fit a jigsaw puzzle. One could employ Shepard-style mental rotation to see, without motor-actional movement, whether the piece fits (Shepard and Metzler 1971; Cooper and Shepard 1973). The Shepard literature is relatively well known by philosophers and was remarkable in its time, with a precise linear relationship between the time taken for match/no-match decisions and the angle of displacement in the matching task (between the comparison stimulus figure and the standard stimulus figure). For instance, some participants, with some stimuli, reliably demonstrated a mental rotation rate of  $50^\circ$  per second. A dispute occurred between those who saw these findings as vindicating the claim that we may think in mental images (and they did vindicate that position) versus those who saw them as compatible with some contrived neo-propositional explanation; but even the latter saw such tasks as involving mental *representations* (albeit propositional representations) then *mental manipulation in the service of problem solving* – here, pattern matching. These data are simply conclusive, and not in dispute. They directly demonstrate the capacity for mental agency as well as anything: in the service of problem solving, we *act to rotate the figure in our mind's eye* (WM tasking offline visual cognition). Making this point is not, however, the purpose of this example.

Note, rather, that one could perfectly well have *physically* rotated the jigsaw piece to solve the pattern-matching problem. And one could perfectly well have *physically* manipulated the joystick or keyboard of the computer game to rotate the Tetris blocks. And one could perfectly well have physically tilted one's head – say, by  $25^\circ$ , shaving half a second of time off the pattern match, reducing thereby the cognitive load of this with any concurrent task. One may physically rotate a map or mentally rotate it or tilt one's head or reorientate one's body in the street – or do a little of each. These examples are repeated *passim* in the situated and embodied cognition literature and are in no sense original to the present work. Motor-actional agency interpenetrates mental-actional agency. There are perception–action *cycles*, and whether the executive actions subserving the cognitive–perceptual task in question are mental (via this route or that) or rather take a detour into the physical environment (via this route or that) is a matter of little importance – these routes are functionally 'interchangeable' (Luria 1979) or, in the modern parlance, 'degenerate' – where 'degeneracy is the ability of elements that are structurally different to perform the same function or yield the

same output' (Edelman and Gally 2001: 13763). This renders deeply problematic any stance that cedes a motor-actional voluntarism fit to subserve an ethical deontologism, yet denies a mental-actional voluntarism fit to subserve an epistemological deontologism – an 'ethics of belief'. Consider, for example, an agent's culpable failure or praiseworthy success in discharging her duty to *find her way* using said map. Motor-actional agency interpenetrates mental-actional agency.

The central idea of the previous section, and a central theme of this chapter as a whole, is that *efferent* cognitive activity massively interpenetrates the putatively *afferent* apprehension of truths about the world. This is something Reid, or indeed any anti-empiricist philosopher from the early modern period, would understand well enough – when we get to the frontal lobes, we are (absolutely) talking about the *active powers of man*. As we ascend away from the 20–25% of the cortex that can be described as sensory or motor, there ceases to be an abrupt distinction between these notions. (Famously, by the time we get to PFC, there are no primary sensory inputs at all.) In higher-level, characteristically human thought, the efferent interpenetrates the afferent: thought (reasoning, problem-solving, judgement) is activity; activity in the service of knowledge and understanding – indeed, activity simply for its own sake. Here, in this section, we note that the motor-efferent interpenetrates the mental-efferent – with each potentially in the service of inquiry. For the example we have been working with, sometimes such action is mental rotation, sometimes head-tilting, sometimes manipulation of a joystick or physical rotation of a jigsaw piece.

Of great, though not exclusive importance here is the role of eye movements – particularly those subserved by the frontal eyefields. One thinks, to some extent, *in and with* eye movement. Neuroscientists sometimes include the frontal eyefields in the PFC, sometimes not (Miller and Cohen 2001), but they have very tight afferent and efferent connections with both motor structures and the DLPFC – with the DLPFC being the classic epicentre of WM, fluid IQ and so much EF and higher cognition generally. The frontal eyefields don't just move our eyes – they *move our thoughts*.<sup>20</sup> We move our eyes (online or offline) to move our thoughts; and we move our thoughts to move our eyes. Thus, Ballard et al. (1997) claim that changing gaze is analogous to changing the memory reference in a silicon computer. They claim that the brain creates its programs so as to minimize the amount of WM that is required, and that eye motions are here recruited to place a new piece of information into memory. Wilson claims that 'eye movements ... allow perceivers to exploit the world as a kind of external storage device' (Wilson 2004: 176–7 – in Clark 2008: 14). Recent work on high-level



cognition vindicates the importance of eye movements in thought, but extends it well beyond storage. Duncker's (1945) famous 'radiation problem' illustrates all the features of the Gestaltists' attractive approach to higher-order cognition in general and problem-solving in particular. It is a difficult, genuinely high-level problem. It involves top-down cognition, insight and 'analogical transfer'. (How do you use a radiation beam to kill a tumour without killing healthy tissue around the tumour?) Familiarly, participants given an analogical problem (rebels storming a fortress) are more likely to solve the target problem (Gick and Holyoak 1980, 1983). But recent research shows that participants who produce certain (voluntary) eye movements in response to a visual representation of the problem are more likely to solve the radiation problem (Grant and Spivey 2003). And this finding is not merely correlational; impressively controlled experimental manipulation shows that bringing about the right kind of eye movements in participants attempting to solve this problem leads, causally, to greater success: '[I]t is now clear that not only do eye movements reflect what we are thinking, they can also influence how we think' (Thomas and Lleras 2007: 668).

Subsequently, these latter authors went beyond eye movements to investigate the classic Maier (1931) two-string Gestaltist problem (described earlier, in Section 4.3 and Endnote 12). In poorly controlled but deservedly famous experimentation, Maier's subjects were more likely to solve the (difficult) problem of tying two strings together, one of which was beyond reach, when cued to the solution (using a pair of pliers as a pendulum weight) by seeing the experimenter brush against one string to send it swinging. However, in cleverly controlled experimentation, Thomas and Lleras (2009) gave participants (who had been deceived that this was an investigation of the effects of exercise and blood oxygenation on problem solving) periodic breaks involving either the exercise of swinging their arms (cueing the 'pendulum' role for the string + pliers that would solve the problem) or of stretching their arms (a cue-lure in the wrong direction). The former (arm-swing) group were greatly more likely to solve the problem. We may see Maier's cue as an *afferent* (passive, sensory) cue – albeit one that assisted switching of the *Gestalt* (the (mentally, not motor) active, top-down knowledge structure, the mental set, the schema) to which the participants assimilated the elements in the problem space. Thomas and Lleras's cue may be seen as an *efferent* (motor as well as mentally active) cue to effect this same switching. Such work shows that 'body movements can actually lead participants toward complex higher order thoughts that they would have been less likely to arrive at otherwise' (Thomas and Lleras 2009: 722). They discuss, speculatively, candidate mechanisms in that paper:

[I]f we assume that spatial working memory is invested both in the planning of movements and in reasoning about problems that are inherently spatial (as the two-string problem is), one can easily imagine that residual activation of location and action plans that a person uses to execute movements might interact with the spatial representations he or she uses to think about the spatial problem. Although the participants are consciously aware of the directed movements that they execute and the spatial representations of the two-string problem space, they do not necessarily have to be aware of the manner in which the former shape the latter. The contents of working memory are conscious, but the route by which they enter working memory need not be. (Thomas and Lleras 2009: 722)

The latter authors see their work as falling within the embodied cognition tradition, but offline neurocognitive activity (e.g. our mental rotation) may serve just as well to facilitate problem solving: the point is, *activity* (whether motor or mental or both) leads to *problem solving*. Note that these Gestalt problems are genuinely high-level cognitive tasks – far more sophisticated intellectual accomplishments than performing the simple perceptual or recognitional or memory tasks with which so much normative epistemology has concerned itself. Cases like this cannot be dismissed or marginalized in the way certain facile examples are set up to be dismissed or marginalized (e.g. as where the agent makes the belief, *p*: *that the light is on*, true by switching it on). Motor movement may mediate the control of self-directed stimulus-independent thought – in the case of the aforementioned research, undoubtedly high-level thought. Thought is mental action; in many cases quite literally mental *action*. One *moves* one's thoughts, one moves one's *mind*, one effects this (mental) process, one *controls* these (mental) *actions*. One thinks in the world with motor and mental actions alike:

Weiner once remarked casually that [a batch of notes and sketches] represented 'a record of [Feynman's] day-to-day work,' and Feynman reacted sharply. 'I actually did the work on the paper,' he said. 'Well,' Weiner said, 'the work was done in your head, but the record of it is still here.' 'No, it's not a *record*, not really. It's *working*. You have to work on paper and this is the paper. Okay?' (Gleick (1993: 409) cited in Clark 2012: 277)<sup>21</sup>

Head movements, eye movements, finger movements, reading, writing, sub-vocalizing, changing one's environment, manipulating one's environment – all these things play a cognitive role. Among other ways, control over such processes may be effected by 'cross-cueing' and 'self-cueing': the kind of motor control used to achieve functional connections outside the body in Gazzaniga's

patients, in Gazzaniga's split-brained rhesus monkeys and *passim* in the clinical (therapeutic–rehabilitative) neuropsychological literature. An internal connection that had been lesioned may be functionally reinstated if an external (motor-actional) connection can be established in its place.

[T]he overall puzzle was why animals that had their brains divided, sometimes far more extensively than ever disconnected in humans, always seemed like they were behaving in an integrated way when it came to carrying out goal-directed behavior. ... Dozens of studies finally revealed one major finding: The animals were engaged in self-cueing. One hemisphere was reading the cues set up by the other. ... Examples of self-cueing come at almost every level of study. (Gazzaniga 2013: 7)

Split-brain monkeys were examined to discover how one disconnected hemisphere could successfully guide the ipsilateral arm/hand toward a discrete object [a grape] in space. The studies revealed that the seeing hemisphere first oriented the whole body toward the object, which cued the nonseeing hemisphere the X and Y coordinates of the object. Subsequently, the appropriate ipsilateral hand posture was formed only when the hand touched the object, thereby cueing the nonseeing hemisphere via touch information. (Gazzaniga 2013: 4)

These connections are not, of course, found only in damaged agents, and do not by any means pertain solely to issues of hemisphericity. In healthy, intact subjects, there are internal-to-internal connections (as with 'cognitive cueing' – efferent control of cognition, but via control of one's *train of thought*) and there are internal-to-external connections, internal-to-external-to-internal connections and all the rest besides. One of the ways we control our cognitive conduct is by motor action, but we also control both our cognitive and motor actions by mental action directly. Our thoughts control our movements and our movements control our thoughts, with some of these movements involving *movement of thought*. There are literally mental actions.<sup>22</sup>

Consider now self-directed, internalized language. One of the basic components of the Baddeley and Hitch (1974) model of WM was the phonological loop. For Atkinson and Schiffrin (1968), a precursor to this (low-level maintenance rehearsal) was the sole method of somehow turning STM into long-term memory – a seventeenth-century model of associative thought ('habit', 'custom', the Thorndike notion of learning as bottom-up experience being 'stamped in'). With Baddeley and Hitch, this becomes one method of maintaining material in WM (and, in conjunction, with the full repertoire of

WM techniques, of working upon said material – as a true EF). But one may vocalize material (a number plate, a telephone number, a name or date) or one may sub-vocalize with a little lip movement, or one may mentally rehearse without any motor movement at all, but with inner speech. Functional magnetic resonance imaging studies indicate that *the same speech areas of the brain are being used in each of these cases* (Ryding et al. 1996, in Barkley 2012) – only in the latter (offline) case, the primary motor areas involving speech are being suppressed. Barkley relates this to the Vygotskian notion of the development of thought through the internalization of speech (cf. Vygotsky 1962: 22). As both Baddeley and Vygotsky note, our self-directed private speech can surely do a lot more than maintain, in a buffer store, someone's car number plate. We think with outer speech and we think with inner speech, with inner speech probably not developmentally, nor perhaps ontologically separable from outer speech or action in the world – it being, as Vygotsky notes, a constituent part of the process of rational activity. Outer or inner speech may mediate thought.

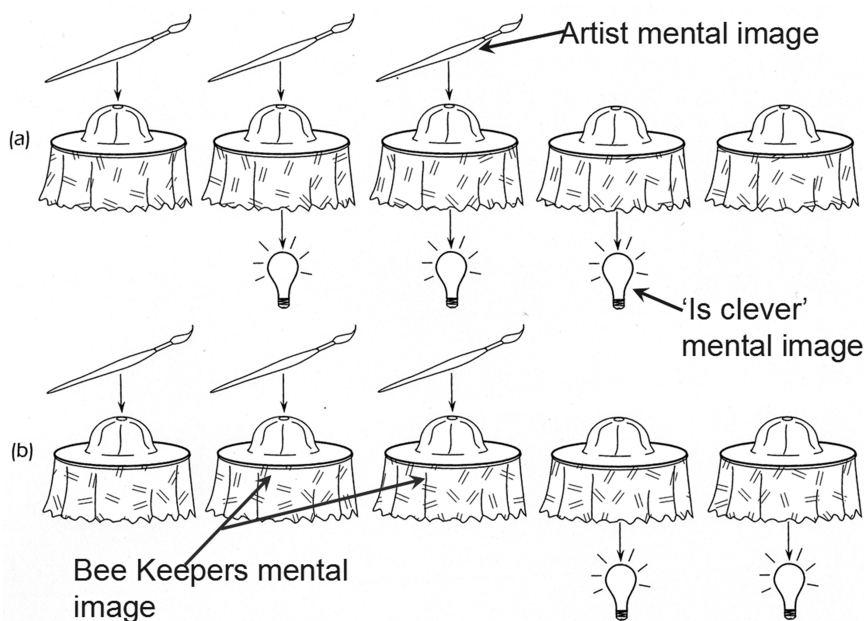
### ***Mental models and physical models***

Philip Johnson-Laird is the figure most associated with the notion of 'mental models'. Were we to situate this notion within a discussion of EF (and I think it naturally belongs there), it would perhaps fit most easily within the notion of WM – though read this claim *modulo* general remarks about holism and promiscuity and reflexivity. There are various issues with the notion of mental models that are of psychological importance (or even of importance in the philosophy of psychology) that nevertheless do not concern me *qua* epistemologist. These include whether mental models commonly or always involve imagery or not (cf. Clement and Falmange 1986); whether they are fundamental or rest upon underlying logical/propositional mechanisms (thus, for example, whether their invocation to explain, say, deductive reasoning carries threats of regress or circularity – Rips 1983); and whether the use of mental models is very widespread in reasoning or relatively less so. Johnson-Laird (2001: 999) claimed, 'Models are the natural way in which the human mind constructs reality, conceives alternatives to it, and searches out the consequences of assumptions.' He might be right, but I would hedge to '*a* natural way'. For my epistemic purposes, it suffices that the notion of mental models is scientifically respectable (is, waiving issues of fundamentality at the level of process explanation, highly likely to be real and psychologically actual at some level and for some/many cases of reasoning) and may be applied fruitfully to explain many cases both of inductive

and deductive reasoning. Now pick a classic example from the psychology of deductive reasoning. A participant is asked to decide what follows, deductively, from premises 1 and 2 of the following:

1. All of the artists are beekeepers.
2. Some of the beekeepers are clever.
3. ...?

Johnson-Laird and Steedman (1978) used exactly this example to support the notion of mental models in syllogistic reasoning. One of the participants in their experiment reported ‘I thought of all the little ... artists in the room and imagined them with bee-keeper’s hats on’ (Johnson-Laird and Steedman 1978: 77). Here you see this with the illustration in Figure 4.3 (courtesy of Sternberg 2009, my arrows added). The top illustration (a) shows the participant’s inductive/belief bias/confirmation bias lure into error; the bottom illustration (b) shows the participant’s discovery (through mental manipulation of the model) that completing the syllogism with the conclusion ‘some of the artists are clever’ will in fact be invalid (Figure 4.3).

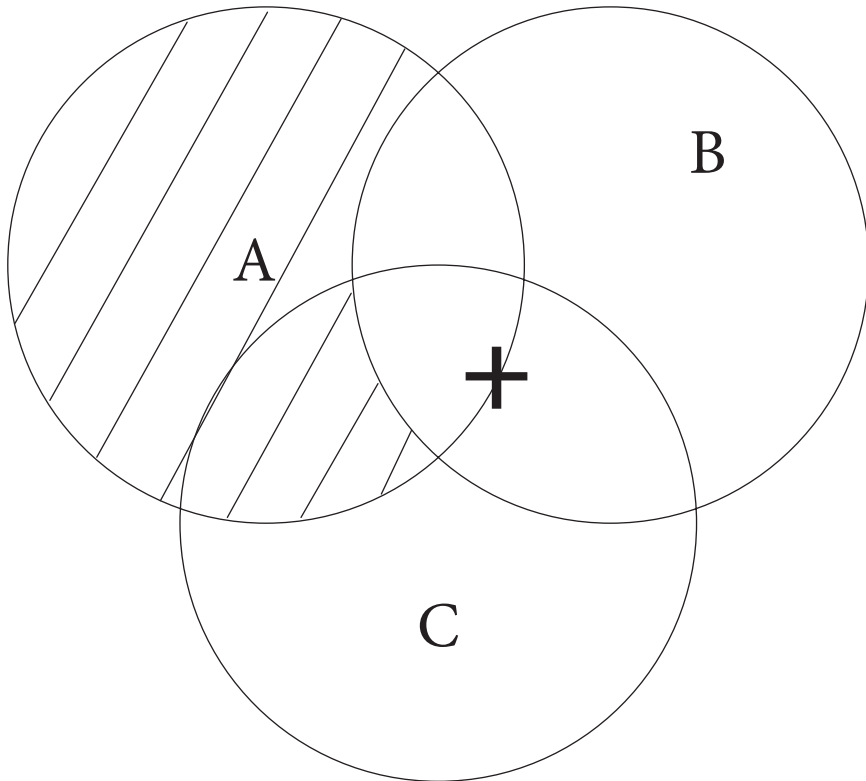


**Figure 4.3** Syllogistic reasoning via mental models (Johnson-Laird and Steedman 1978, as illustrated in Sternberg 2009)

Source: Sternberg. *Cognitive Psychology*, 5E. © 2009 South-Western, a part of Cengage Learning, Inc. Reproduced by permission. [www.cengage.com/permissions](http://www.cengage.com/permissions)

Notice, however, that one did not need to make this discovery through construction and manipulation of a *mental* model. Many of us were, as undergraduates, taught, prefatory to embarking upon a first course in predicate logic, how to prove Aristotelian syllogisms through physically drawn models – with Venn diagrams. Here, in Figure 4.4 is our earlier syllogism. A, B and C are the three sets of artists, beekeepers and clever people, respectively. ‘All’ is indicated by shading out the region A-B. ‘Some’ is indicated with a cross (indicating an element) in  $B \cap C$ , which cross is placed on the line, indicating that there is no information contained in the premises as to whether this element is a member of the set A. With the diagram disclosing no information as to whether this element is in the set of artists, the claim that ‘some of the artists are clever’ is shown to be invalid.

What is the fundamental epistemic difference between this and the former proof? Each establishes something with apodeictic certainty by construction of a model. One is a model drawn in chalk or graphite in the physical world, one



**Figure 4.4** Syllogistic reasoning via Venn diagram

a mental model. It could be the same model if you like – an agent well versed in the pencil and paper version of a Venn proof constructing said proof in her head; or (conversely) a person establishing the proof by drawing on paper the picture illustrated in Figure 4.3.

In each of these cases, we have WM operating through the manipulation of elements in a visuospatial workspace. Fleshing the example out with the (optional) ontological commitments of the Baddeley and Hitch (1974) model: in one case, this is *mental* manipulation by the *central executive* operating on the contents of the *visuospatial scratchpad*; in the other case, it is *mental and physical* manipulation by the *central executive* operating on a *physical notepad* – and *mutatis mutandis* should you prefer the ontological commitments of some other model. Note, importantly, that on many/most neuroscientific concepts of EF, the frontal lobes do much EF through activating and reactivating and directing cognitive operations that occur in more posterior cortical structures – and some lower-brain structures, and (here) some physical structures in the world. The ‘central executive’ is *directing, activating, reactivating* and *manipulating* visual imagery mediated by the same structures as view the physically manipulated/constructed diagram, only (here) doing so offline. I can prove this syllogism to be invalid with the construction then manipulation of a physical model or a mental model or a little of both. In each case, I am using cultural scaffolding that is outside of the cranium. (I didn’t discover syllogistic logic or naïve set theory or model theory *ab initio*, nor learn for myself how to manufacture and deploy paper and pencil, nor give myself the years of training needed to form a decontextualized mind: Cole and Scribner 1974; Luria 1976; Vygotsky 1978; Lockie 2016a,b.) Metacognition indeed may become sufficiently anterior as to burst out of the front of the forehead. But metacognition (EF generally) also tasks and governs lower cortical (and not even cortical) functions. In this and many cases, it does so through (culturally mediated) model building, with the issue as to whether this model is manipulated partly in a physical or only a mental workspace being of relatively slight psychological significance and of no obvious epistemological significance at all.

As for this specifically imagistic conception of a mental (/physical) model – that is, in a *visuospatial* workspace, so *eo ipso* for other, non-visual imagery cases. We may construct imagistic models in other sensory modalities, or in natural language, or other neo-linguistic media like non-visually mediated mathematics, or any other deemed to be epistemically respectable medium. You take the medium (natural language, say) that avowedly mediates a specific case of epistemically respectable first-order cognition. You have storage and executive

capabilities – the ability to work upon and *play* (Barkley) in this workspace. You perform hypothetical, decoupled thinking in this workspace. Thus do you make discoveries. Take the case of model construction in the medium of natural language. Some of these models may be in spoken language, some written, some in inner speech. You have real arguments with extant protagonists, you have written exchanges, you have spoken exchanges, you read, you annotate, you jot down notes, these become a draft, you have *l'esprit d'escalier* (you construct exchanges in your head with an imaginary interlocutor – exchanges that you win); thus do you construct your arguments, you discover and defend what you take to be the case. You reason (here: linguistically) offline and online and a little in-between, and you reason with the same faculties in each case. You use your intellectual freedom thus to present yourself with your chosen problems, construct models appropriate to address these problems, 'play' in that decoupled problem space and, if you are lucky, perhaps even solve said problems, or at least reason yourself until profoundly satisfied, by your own deepest standards, as to what you take to be the case: achieving deontic epistemic justification (if you like, becoming Foley rational) from actively seeking after the truth with all the resources at your disposal, all the *active powers* you possess.

I do not see how these connected points – regarding the *cognitive* interpenetration of the motor-actional and mental-actional – can be conceded (and they must be conceded), yet the abrupt distinction between motor-actional voluntarism and mental-actional involuntarism be maintained. Grant me freedom of motor action and a good deal of freedom of mental action comes, well, for free.

## 4.7 Two-process theory

Noted already is the tendency among numbers of philosophers to conceive of cognition most generally as unconscious, mandatory, automatic, autonomous, non-rational, fast, parallel, etc. In the (optional) terminology of the two-process literature, these are Process/Type/System 1 characteristics. In that literature, these characteristics are precisely contrasted with higher-order, rational, conscious, slow, serial, decoupled, hypothetical, controlled, metacognitive, etc., cognition (Process/Type/System 2 characteristics – the more distinctively human psychological abilities). The ontological status of this distinction is an important, disputed, issue in the philosophy of psychology; and whether it is a well-founded distinction at all is a major, live debate in theoretical psychology. Two theorists



who have separately done more than most to champion this distinction are Jonathan Evans and Keith Stanovich. Both are essentialists about this distinction. Evans takes as definitional for Type 2 cognition that this be under the control of WM.<sup>23</sup> In this he diverges from his recent collaborator Stanovich, who emphasizes the human capacity for cognitive ‘decoupling’ – though in their first ever jointly authored paper (Evans and Stanovich 2013), they claim that their core criteria are compatible. In the following useful table from that paper, listed here as Table 4.1, they list their (separate) definitional criteria in italics, followed by ‘correlates’ (which, in contrast, non-essentialists might choose to use disjunctively, or explicatively, or constitutively – say, to defend the distinction as a cluster concept).

One might see the left-hand column of Table 4.1 as a scientifically far more developed version of the right-hand column of the little ‘toy’ table (Table 3.1) in

**Table 4.1** Two systems/types/minds (from Evans and Stanovich 2013)<sup>24</sup>

<b>Type 1 process (intuitive)</b>	<b>Type 2 process (reflective)</b>
Defining features	
<i>Does not require working memory</i> <i>Autonomous</i>	<i>Requires working memory; cognitive decoupling; mental simulation</i>
Typical Correlates	
Fast	Slow
High capacity	Capacity limited
Parallel	Serial
Nonconscious	Conscious
Biased responses	Normative responses
Contextualized	Abstract
Automatic	Controlled
Associative	Rule-based
Experience-based decision making	Consequential decision making
Independent of cognitive ability	Correlated with cognitive ability
System 1 (old mind)	System 2 (new mind)
Evolved early	Evolved late
Similar to animal cognition	Distinctively human
Implicit knowledge	Explicit knowledge
Basic emotions	Complex emotions
Note. Italicized attributes are the proposed defining characteristics in the current article. Authors proposing two systems include the features attributed to Type 1 and 2 processing but may also include the additional features named.	

Source: (Evans and Stanovich 2013). Copyright Sage 2013.

the preceding chapter: the mandatory lower-brain (or at best more-posterior-cortex) ‘forced’ cognition of the involuntarist epistemologists. Notice now how Evans and Stanovich’s right-hand column *itemizes a list of characteristically executive cognition*. In fact, I suggest that the best sense we can make of a principled criterion for Type 2 cognition is that *Type 2 processing just is executive processing*<sup>25</sup> – with Evans’s WM and Stanovich’s decoupled processing each then seen as attempts to capture the essence of such processing in terms of a unitary, core feature of EF. Given (as we have seen) how complex and disputed work on executive functioning is, using this concept to capture the essence of the two-process distinction may be criticized as an *obscurum per obscurius*, but I think there has been remarkable progress in understanding EF – some of which has been summarized in this chapter. In any event, I think the connection mooted is likely true, and here (guardedly) suggest we identify Type 2 processing with executive functioning, the better thereby to make a critical point.

It is an empirical truth that we, as a species, possess the extraordinary, distinctively human EFs articulated in the foregoing, and captured, I would suggest, in the right-hand column of Evans and Stanovich’s table. In light of this truth, a standing challenge for my involuntarist opponents is to articulate how the claim that our cognition is forced may be qualified to come out as true – then to work forward from this to state what normatively (epistemically) a statement so qualified is capable of warranting. These, I contend, are *hard questions* for my opponents; but in light of the weight of empirical evidence cited, I would suggest that the best case that could be mustered on their behalf would involve noting that certain ‘old mind’ (Type 1) aspects of cognition are indeed ‘involuntary’ in the sense at issue – hence not deontically assessable. Indeed, in a sense, such cognition – at least considered in the absence of other (Type 2, executive) characteristics of the cognitive agent who owns it or tasked it – may appear not to be *epistemically* assessable at all (that is, assessable in terms of normative epistemology, as opposed to psychologically assessable, say, for speed or accuracy). In any event, such cognition does not obviously appear apt for assessment in terms of its epistemic rationality, or lack thereof. But the question in these terms then becomes whether such cognition may be delineated as a class without appeal to *just the notion at issue* (you choose: ‘involuntary’, ‘not consciously accessible’, ‘encapsulated’, ‘automatic’, ‘not involving intellectually normative responding’, ‘mandatory’, etc.), and this appears to be a *decidedly sharp* question for these ‘involuntarists’.

Since, however, no sensible epistemic deontologist would seek to deontologically assess these (Type 1) kinds of mental operations or outputs (at

least, absent the Type 2 cognition that put them in process, or with which they must mesh and integrate at the agential level), the critical, epistemic point to noting the ‘forced’ (mandatory) nature of such ‘old mind’ cognition is extremely hard to fathom – and this quite apart from the aforementioned challenge of articulating any such ‘forced beliefs’ claim at all without viciously begging the question. We deontically assess epistemic agents, their thought processes, their arguments, their beliefs, on the basis of their control of their ‘Type 2’, distinctively human capabilities; or at least with their untyped cognition more generally, yet where this be bound up with such capabilities: with things that the epistemic agent may be held responsible for.

#### 4.8 The bias argument

The EFs indeed precisely *bias* our lower-brain belief-forming mechanisms (‘bias’ is a term repeated *passim* in the neuro-cognitive literature<sup>26</sup>). But this doesn’t ‘pervert’ (Kornblith) our cognition – it is *essential* to high-level cognition. In their influential review of PFC function, Miller and Cohen note the following:

the PFC serves a specific function in cognitive control: the active maintenance of patterns of activity that represent goals and the means to achieve them. They provide bias signals throughout much of the rest of the brain, affecting not only visual processes but also other sensory modalities, as well as systems responsible for response execution, memory retrieval, emotional evaluation, etc. The aggregate effect of these bias signals is to guide the flow of neural activity along pathways that establish the proper mappings between inputs, internal states, and outputs needed to perform a given task. This is especially important whenever stimuli are ambiguous (i.e. they activate more than one input representation), or when multiple responses are possible<sup>27</sup> and the task-appropriate response must compete with stronger alternatives. From this perspective, the constellation of PFC biases – which resolves competition, guides activity along appropriate pathways, and establishes the mappings needed to perform the task – can be viewed as the neural implementation of attentional templates, rules, or goals, depending on the target of their biasing influence. (Miller and Cohen 2001: 171)

‘Bias’ here and in this literature more generally is, I believe, used in a somewhat ambiguous fashion. Narrowly, it can mean bias in an engineering sense (in terms of various theories of biased competition and contrast and signal gain, of the kind commonly employed in theories of executive attention (cf. Wu 2014) and derived ultimately from electronic engineering). More broadly, it means that EFs prevent

cognition according to prepotent, merely associative pathways from prevailing, but by means of various mechanisms (perhaps not even/only *mechanisms*) as yet poorly understood – of which bias in the narrower sense would be but one. I am happy to live with any such ambiguity and do not see that it affects my arguments. The point is, EF ensures that the prepotent pathway does not always prevail. It ensures that weaker, task-relevant pathways are chosen over stronger but task-irrelevant pathways. The posterior cortex mediates more automatic, less controlled processing. The job of more anterior, executive processing is to *bias* this processing (n.b. and lower-brain, non-cortical processing). That's what good cognition is:

According to our current view, common EF is about one's ability to actively maintain task goals and goal-related information and use this information to effectively bias lower-level processing.<sup>28</sup> This basic ability is necessary for all three EFs [inhibition, WM and flexibility]. (Miyake and Friedman 2012: 11)

This lower-level processing, sans control, sans 'bias' (seen by the doxastic involuntarists as a *contaminant* of good cognition), is said involuntarists' whole, crippled, conception of cognition – their paradigm of afferent, passive, associative, received knowledge. Only it's not so passive. We have not simply the mediate powers of functional control already emphasized, but directly active cognitive powers. What would happen if we didn't have this top-down avowed biasing of our otherwise passive, afferent, sensory, mnemonic, etc., systems? Tragically, from the clinical literature, we know only too well:

The biasing influence of PFC feedback signals on sensory systems may mediate its role in directing attention ... signals to the motor system may be responsible for response selection and inhibitory control ... and signals to intermediate systems may support short-term (or working) memory ... and guide retrieval from long-term memory. ... Without the PFC, the most frequently used (and thus best established) neural pathways would predominate or, where these don't exist, behavior would be haphazard. Such impulsive, inappropriate, or disorganized behavior is a hallmark of PFC dysfunction in humans. (Miller and Cohen 2001: 172–3)

Bias doesn't 'contaminate' veridical cognition; it is a prerequisite of high-level, human, rational cognition. Without bias, the strongest ('prepotent') pathway always prevails – you have no control, you cease to be a high-level cognitive agent at all, you become a Skinnerian 'locus of variables', an associative adding machine – or the randomized (classic *Mind* argument) alternative (Miller and Cohen's 'haphazard', 'impulsive', 'inappropriate' or 'disorganized' behaviour).

Unimpaired agents *bias* their lower-brain, purely sensory systems to effect good, *controlled* cognition. We can relate this to the classic Hobbes–Hume–Hobart dilemma in the free will debates – which will shape so much of Part 2 of this book, and especially its last three chapters. ‘Unbiased’ cognition is (at best) the rigid, inflexible, associative, strongest-motive cognition of the kind found in lower animals and in such modular/mandatory/automatic processes as we share with said animals – the determinist horn of that dilemma. At worst it is random and haphazard – Jonathan Edwards’s (2007: 151) ‘smoke that is driven by the wind’ – the indeterminist (randomness, chance) horn of that dilemma. Both are found in frontal lobe patients. Neither is a model of good (higher-order, characteristically human) cognition, which is precisely why Hobart’s dilemma is a false dilemma – just as the agent-causationists always maintained (Reid 1969: Essay 4, Ch. 4, #3, #5; also Nozick 1981: 294–306; McCall and Lowe 2005: 686ff; Steward 2012: 144ff). Frontal lobe patients and lower animals without EF are not blameworthy or unjustified cognizers precisely because they can’t *control their cognition*; but in virtue of being thus ‘unbiased’ they are hardly thereby epistemically justified agents. They are not epistemic *agents* at all.

A kind of parallel to this finding was already understood by psychologists in the distinct area of knowledge (/memory) research by the great Bartlett–Neisser constructivist tradition (Bartlett 1932; Neisser 1967). All high-level, characteristically human cognition is biased, top-down. Bias doesn’t ‘pervert’ (‘contaminate’) an otherwise pristine pathway; ‘preservation’ theories of cognition, whether as regards reasoning, memory or knowledge, are (at least in the general case and as regards *higher* cognition) both entirely indefensible and achingly naïve – and this as regards the scientific psychology and the normative epistemology alike. Without these abilities (‘biases’), we aren’t human cognizers – we aren’t even mammals (cf. MacDonald’s 2008 ‘hot sand’ example, cited above).

#### 4.9 The executive functions and regress

There is an obvious, distinctively philosophical problem – of regress. We want to know how normative epistemology is possible: how can we *control our thoughts*? I have taken my answer from the psychologists (and, in inchoate form, from common sense): *through executive functioning*. I contend that this extended, empirically informed answer is simply a psychological truth (hardly a counterintuitive one either) and ought never to have been called into question,

least of all with arguments some of which purport to represent the worldview of modern psychological science.

But the question arises: does the exercise of these and similar EFs constitute *action*? Comes the answer: yes – these functions, when exercised, constitute actions; in many cases they are more paradigmatically actional than most non-executive and specifically motor actions. But then what *causes* (/brings about) these EF actions? Are they determined (triggered) involuntarily (e.g. mechanistically, by the stimulus: internal or external)? Or are they caused by, say, *intentions* in turn? You can see where this is going. To pick one horn of this dilemma: if the proximal determinant of an EF is a mechanistic sub-agential process, what is the net gain *philosophically* of noting the role of the function that is then brought about as itself a contributor to what is held to be the non-automatic, controlled modification of cognition?<sup>29</sup> But consider the other horn of the dilemma, where an EF action is caused by, say, an agential-level ‘intention’: the same dilemma presents itself regarding this ‘intention’, and the danger of regress is obvious.

Here, a few points are worth registering. One point is that the epistemic problem we were confronted with (‘doxastic involuntarism’) was *not* meant to be the problem of free will per se (say, in the teeth of Hobart’s Dilemma or Kim’s Exclusion Argument or Strawson’s Basic Argument or some other fundamental sceptical challenge). It was a prescribed, much more specific problem: granted that *motor actions* are indeed actional, hence ethically deontically appraisable, nevertheless, it was asserted that our beliefs are forced. Sometimes this was taken to be an armchair truth, sometimes the clear verdict of the cognitive sciences, sometimes the outcome of a very seventeenth-century argument which said that in order to be *epistemic* at all (a putatively veridical deliverance of our faculties), our cognitive faculties cannot be *free* (the product of our ‘desires’) as this would *bias, contaminate, pervert* their normative epistemic status. But we have comprehensively dealt with these objections, and that was the task we were set by the involuntarist opponents of a deontological conception of epistemology. The first two are egregiously false and/or contrived, the third involves a remarkably *insouciant* conception of good cognitive functioning on the model of ‘uncontaminated’ cognition: the ‘preservation’ theories so beloved of armchair psychologists and so relentlessly, utterly, refuted by the last hundred years of empirical psychology – to say nothing of the Kant/Reid philosophical traditions prior to that. Little higher cognition is to be understood (factually) on the model of a preservation of afferent signal accuracy, of ‘labelled lines’, in from the sense organs to the centre – not as much as is sometimes believed

even of perceptual or mnemonic cognition, and very little of the *active powers* of man.

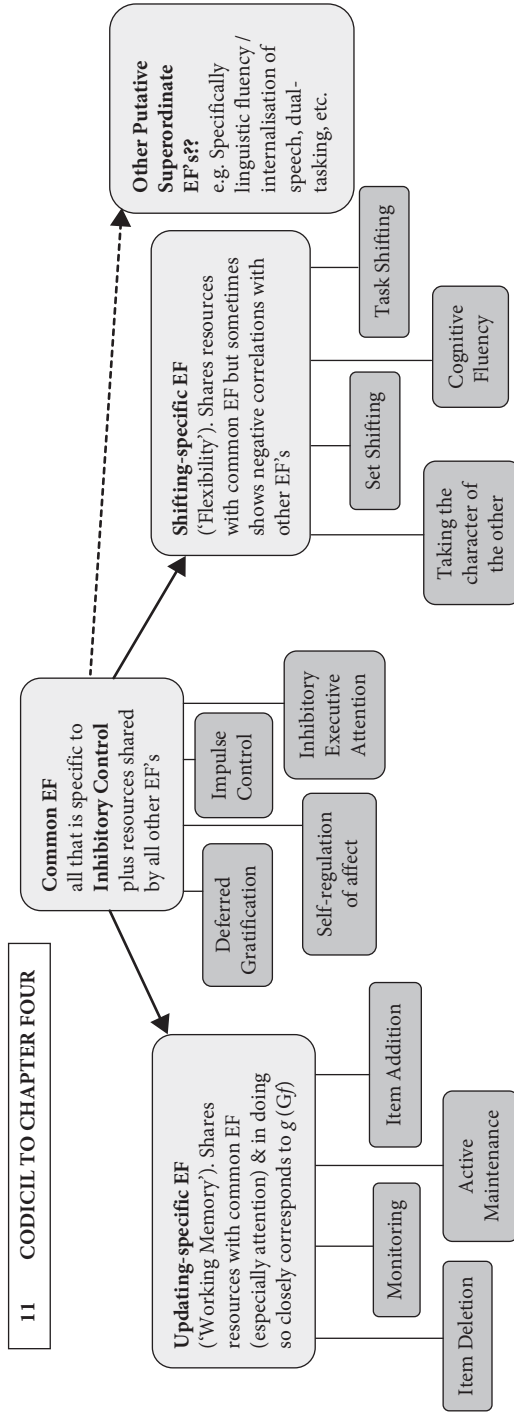
Were epistemic deontologism indefensible because no action (motor *or* mental) could really be under our control in a way that makes responsibility (any responsibility) possible, that would be another argument, one far more sweeping than the specific epistemological debates about voluntarism. I shall address issues of free will and responsibility more generally in Part 2; but these debates are quite different from (are far more general than) the epistemic debate. I take it that it would be a sign that this latter debate was well won if the specifically epistemic involuntarist had to change the subject and generalize matters so drastically.

Now consider the regress objection to epistemic freedom in its own right. It is exceedingly swift, and I think by no means convincing. Firstly, note the danger of mereological fallacy in claiming that since many (not all) of the parts that comprise an EF are automatic and uncontrolled, therefore the whole cannot be controlled. The parts that make up a cell are not alive, but the cell is. Secondly, note the neo-behaviouristic, neo-Wittgensteinian philosophical tendency of assuming that since a regress potentially beckons, therefore the first step should not be taken. The EFs are clearly empirically actual and a central, core aspect of what our epistemic agency consists in. Philosophical worries about a potential regress cannot alter that fact: it is an unassailable psychological truth; and one does not build a satisfactory philosophical account by ignoring, as an important descriptive starting point, a clear empirical truth simply on the grounds that it is not 'ultimate' enough for one's longer-term explanatory (/reductionist) tastes.<sup>30</sup> Thirdly, note that while some aspects of the parts that comprise an EF are automatic and uncontrolled, some (many) are themselves controlled. Fourthly, note that an alternative to a regress may be a circularity: the EFs continually, reflexively, systemically affect and control each other and lower-brain processes and actions in the world. This begins in infancy and ends in death. There is a pervasive functionalist, systems-level, self-organizing, diachronic developmental process (these are *processes* after all). Fifthly, note that these functions are not simply in the head: they are situated in the body and in the world and interpenetrate our cultural–cognitive achievements in open-ended, potentially unlimited systems of control: EF permits the development of culture, and culture radically changes the nature of EF. Sixthly, note the possibility of emergentism (highly compatible with much of the above – e.g. points 1 and 4, and points in the text and Endnote 14 regarding token versus type proximal determination of mental action – an emergentism itself compatible

with an agent-causal (or, I shall say in Chapter 9, a self-determinist) conception of matters). Finally, note that intellectual freedom is plausibly an agential-level predicate rather more than it is a process-level one (though it may be both). The EFs, as noted, control and *bias* many other EFs and many lower-brain functions – which feed back to them and into the world (cf. Miller and Cohen on the EFs, Bartlett, etc.: the constructivist tradition in top-down cognitive psychology *passim*). I hardly think that these brief remarks are sufficient to set aside the problem of free will as it confronts us in normative epistemology in the longer term, but they throw back the onus upon the involuntarist to defend his specific regress objections, meant, of themselves, to swiftly dismiss, without even addressing, the constructive psychological points and clear descriptive/scientific truths that I have raised in counter to that position above.

This chapter and the last have addressed, with scientific evidence and constructive argument, how the intellectual freedom necessary to defend a deontic conception of internalism could be actual; and why we have strong reasons to believe these powers are actual. The next chapter is wholly different in tenor; eschewing any scientific considerations to advance a negative, transcendental argument: that at some level and to some extent we *must* embrace the deontic conception of internalism – that in this, we have no choice.





**Figure 4.5** Illustrative example of a hierarchical model of executive functioning (not endorsed by the present author). Miyake & Friedman's (2012) factor-structure for the three highest superordinate EF's is given (light shading), together with the subordinate EF's they mention under 'Updating'. Merely illustrative subordinate factors are included under other superordinate EF's (darker shading)

- who are impaired prior to that (congenitally, say, or through traumatic brain injury in early childhood). (This is true of socioaffective EF impairment also.) EF involves both hot and cold cognition, each of which is significantly mediated by the frontal lobes.
- 7 Later versions of Shallice's theory became considerably more complex.
  - 8 **Letter-memory task:** update WM by remembering the last few words/letters of an ongoing list. **N-back task:** you are presented with a series of stimuli, one every two or three seconds, and must decide if the present stimulus is the same as one presented *n*-items back in the list. **Flanker task:** you must indicate the direction of a central fixation target arrow using a keypad, sometimes when its flanking (distracting) arrows are pointing in the same direction (congruent) and sometimes when they are pointing in a different direction (incongruent – thereby needing inhibition). **Go/no-go task:** you press a button on each presentation of a letter stimulus as fast as you can – except when the letter is an 'x' (when you must inhibit the prepotent response). **Trail-making test:** connect the circles containing letters and numbers in order, only alternating between connecting by number and by letter (thereby having to 'shift'). **The Wisconsin Card Sort Test (WCST):** each card in the deck has one, two, three or four of the same shape on it; these shapes are one of four different kinds (triangles, stars, etc.) and in one of four colours. You have to guess whether you should sort by number, colour or shape and are told only 'right' or 'wrong'. A while after establishing the correct rule, it changes: how quickly can you detect this and change your sorting principle? It is only fair to note that there are longstanding doubts about the psychometric properties (especially reliability) of tests like these; and there are issues as to their short timescale and the fact that they exclusively measure cold cognition. EF is commonly measured by tests like this, but, constitutively, it is a lot more than the cognitive processes subserving only these things.
  - 9 Immediately applied gratification delay (resisting the cream cake) would be an example of *impulse control* (an inhibitory EF) – or perhaps 'willpower' to the extent this is something different. Prosthetically/functionally applied gratification delay (not shopping when hungry, not stocking your refrigerator with cream cakes, avoiding stimulus cues by covering tomorrow's tea party's cream cakes from sight in the refrigerator) would be an example of metapsychological regulation; also in this case a (higher-level) inhibitory EF to some extent, but one involving other EFs like WM and theory of mind and umbrella-term EF's like 'planning' – besides, crucially, operating over a much longer timescale.

## Chapter 4

- 1 Quite a few (neuro)psychologists, especially those influenced by factor analytic approaches, do separate out something like three superordinate factors in this

fashion. It originated with Friedman et al. (2008) and Miyake and Friedman (2012). I have illustrated a version of their model (similar to that which is endorsed by Diamond) in Figure 4.5. Some distinguish a fourth: specifically, verbal fluency (sometimes ‘self-directed speech’). Best et al. (2009) have as a fourth ‘planning’ (to include ‘problem solving’). Others distinguish ‘dual tasking’ or ‘interference control’ (Willcutt et al. 2005). However, many don’t take such a ‘superordinate grouping’ approach, and the debates about whether such superordinate factors are psychologically real or correlational artefacts are well worked in, for example, the individual differences literature. (Exploratory factor analysis, for all its complexity, is only a descriptive statistic, and the factors extracted are very strongly determined by several arbitrary matters under the control of the researcher – here, notably, the fact that the inputs to this analysis are exclusively short-duration, cold-cognitive tasks.) Note that for any of these superordinate factors there is the problem that they may be too broad and inclusive to be seen as more than expository grouping variables, and too disjunctive at the process level, which is anyway little understood. Also, as already emphasized, finding one putatively orthogonal factor (‘WM’) requiring or logically presupposing the actions of another (‘inhibition’) is a phenomenon that is simply ubiquitous. In contrast to Diamond’s confident assertion, Barkley (2012: 8) strongly repudiates any claim of consensus here, talking of the ‘dog’s breakfast of constructs that EF has become’.

- 2 Some researchers prefer to see this as simply a tough measurement problem (‘task impurity’: the problem that a given operationalization doesn’t measure only the one EF, or even only EFs rather than lower-brain functions) – a problem to be overcome either with more careful experimental isolation of one’s dependent variable or with greater statistical sophistication (e.g. the use of latent variable analysis; Friedman et al. 2008; Miyake and Friedman 2012). But the low *process-behaviour correspondence* for EFs (that similar behaviours may have quite different causes) may not be a source of error, (‘task impurity’), being due rather to features that Luria (say) would have immediately understood: the fact that executive functioning is intrinsically holistic, functional, systemic and coordinated. There are lessons to be learnt here from the Bartlettian objections to the Ebbinghausian tradition in memory research: ‘purify’ your DVs too much and you may well throw the baby out with the bath water. ‘Task impurity’ may, then, not be a methodological problem, but a profound functionalist, holistic truth about the *in vivo* higher cognitive functioning of the human mind.
- 3 It is not only an elastic notion, but also (like so many of these EFs) has tendencies to be somewhat homuncular – something Baddeley honestly acknowledged of his ‘central executive’ from the start. I am concerned to identify scientifically respectable, roughly agreed truths about the extent of our psychological powers: descriptively accurate accounts of our abilities that any normative epistemologist

may make appeal to, and with which any acceptable normative epistemology must comport. I do not then feel that such matters, however much they may be of concern to psychology or even to the philosophy of mind, vitiate my deployment of this and cognate notions, nor require much further engagement in this (epistemological) context of inquiry (although homuncularism I engage with further below).

- 4 Using TOTE to illustrate WM is borrowed from Conway et al. (2011), but the driving example is my own and is introduced for three reasons, the first being its ecological validity: weak executive functioning is massively predictive of poor, reckless, accident-prone and inconsiderate driving. The second is that it foregrounds variegation in scope as a key feature of both normal and impaired EF: some of these driving problems are more narrowly cognitive (cold cognitive: Norman 1981; Reason 1992 on field dependence, etc.), some more cognitive–affective, some more personality-mediated or agential. The clinicians emphasize things that the cognitive psychologists insufficiently emphasize: that hot cognition and the socioaffective nearly always interpenetrate cold cognition when it comes to executive functioning. Finally, driving is one of the classic examples used to indicate that what counts as executive functioning is not static: what was at one stage an EF becomes, through processes of automaticity, a more posterior cortical function and eventually a lower brain function. This latter is of importance when we reach the section dealing with two-process issues below.
- 5 This posterior-to-anterior ascent in hierarchical abstraction in rules and reasoning is found in the DMPFC and DLPFC, which have major anterior-to-anterior and posterior-to-posterior connections with each other. There are claims that something like this posterior-to-anterior ascent in abstraction is also found in the ventro-medial PFC (VMPFC; Bechara and Damasio 2005), with the VMPFC occupying a crucial role in mediating social and moral cognition. It is highly likely that these three key areas in the PFC play a major role in deontic reasoning, with there already being evidence that VMPFC-lesioned patients are more utilitarian and less deontic in their judgement of classic (ethical) trolley-car cases (Greene et al. 2001; Greene et al. 2004; Greene 2007; but see Moll and de Oliveira-Souza 2007) – presumably because their judgements are solely informed by consistency and other types of ‘cold cognition’ (as mediated by the DLPFC) rather than also by the ‘hot cognition’ mediated by the VMPFC. If there is a deontic aspect to our learner driver’s invocation of checking procedures, then I would warrant a role for VMPFC. To the extent this remains *reasoning*, I would expect, however, that it involves VMPFC in close systemic interaction with DLPFC.
- 6 Jung and Haier’s ‘parieto-frontal integration theory’: sharing resources across a network of regions predominantly in the parietal lobes and frontal lobes in an integrated fashion – cf. Haier (2017).

- 7 The separation per se isn't (here) claimed to be illicit: insisting sans motivation that the deontologist must be committed to defending voluntarism with respect to the avowedly automatic 'final process' is; while insisting/assuming (sans argument) that the deontologist's actual, expressed defence of deontology of the processes leading up to these somehow doesn't answer to her specifically *epistemic* needs. Elsewhere (in Lockie 2014b and a planned monograph on knowledge) I 'problematize' certain lower-to-higher stage separations of this nature per se (in arguments referring to historic disputes of Bartlett versus Ebbinghaus, and Titchener versus the Gestaltists – the idea that the putatively lower level is unvarnished and prior *data*: prior, that is, to the higher-level, top-down putative 'contamination' of that data).
- 8 I have too little space to address issues pertaining to personal/practical cognition with the depth they deserve, but awareness of self across time ('autonoetic awareness') is a distinctively human, late-developing capacity that involves/integrates with controlled reactivation of memories, planning over variable time-and-distance frames, capacities for 'decoupled' (hypothetical) thought and model building involving self-relevant counterfactuals over time. Can such cognition, involving 'prospection' (Seligman et al. 2013), be deontically appraised? Yes, in many instances. Does such cognition involve agential-level control? Yes, paradigmatically. 'The capacity to imagine a hypothetical future from an experienced past is one of the three most important or foundational EFs. The other two are self-directed attention (self-awareness) and self-restraint ... together they create the human sense of the future' (Barkley 2012: 85). Can we deontically assess an agent (as commendable or remiss) for their *consideration of future consequences* (or lack thereof)? Of course we can, and do. It is even operationalized as a named cognitive–personality variable (Strathman et al. 1994). It is partly mediated, anatomically, by regions that are of importance in deontic reasoning more generally (the VMPFC: Bechara et al. 1994, 2005). A deontic conception of epistemic responsibility is often assimilated to a very 'synchronic' notion of responsibility (e.g. Foley, *passim*); but cognitively, I can deploy the EFs Barkley identifies to travel in time and space, and I may be held responsible for failing to do so.
- 9 'Directing and redirecting executive attention to goal-relevant information may be the primary mechanism by which self-regulatory goals are "shielded" from competing goals or other distractions. ... According to this view, goal shielding is the consequence of sustained attention to a goal or task and provides an indirect or "passive" form of inhibitory control. ... The ensuing mental state in which individuals "zoom in" on the goals they want to achieve may closely correspond to what has been called an implementation or action-orientation mindset in self-regulation research' (Hofmann et al. 2012: 175). Notice the promiscuity/reflexivity motif here: executive attention, WM, inhibitory control, self-regulation,

passive, active ... one explaining the other and then explained by the other in turn. One could seek to clarify a structured path through such accounts, but I think (conceptually, not empirically) that would be unwise. The functional, holistic interpenetration of EFs *in vivo* is, I think, a profound and important truth about the higher cognitive powers of man.

- 10 Will this moment of insight then be defined into our involuntarists' 'final pathway' for non-sensory cognition? I take it any such move would simply represent a determination to find a terminological sop for a substantial philosophical defeat. Note that Heisenberg will work forward (in his controlled, executive cognition) from this insightful discovery also (into further mathematics) and will have developed the mathematical expertise that permitted this moment of insight significantly because of his preparedness to work as doggedly yet flexibly as this for many thousands of hours previously over the years; and (more proximally) for many scores of hours previously on this specific problem. His thought is massively diachronic, is pervasively and systemically the product of many different *perception–action cycles* (Fuster 2013). These cycles, occurring at every level of cognition from the lowest to the highest, involve afferent and efferent processes partly tasked by the agent and partly (as diachronic, developmental and self-organizing *processes*) leading to the development of the agent, ontogenetically and aretaically. No intellectually serious thinker working on these issues doubts that automatic and uncontrolled cognition is bound up with creative, controlled, executive cognition – or that (in unimpaired agents) each is present in a single mind: cf. remarks to follow on two-process theory.
- 11 In scare quotes because creativity as an *agential*-level property is a product of all the EFs and much more besides.
- 12 We deal with an important update to Maier's two-string experiment below. Maier's subjects had been given the problem of tying two strings together that were suspended from a ceiling. The strings were long enough to reach each other if the participant could reach both, but the participant could not stretch to reach the second string whilst still holding the first. A range of objects were in the room that the participants were permitted to use. The correct solution was to tie one string to a pair of pliers then use this as a pendulum to swing this string towards the first until able to grasp both together. This task was difficult for participants because they were *functionally fixed* on the pliers as a gripping tool – not a weight. This joins several other Gestalt cases as a classic investigation of 'insight' problem solving, which in our modern terms mostly concerns *flexibility of thought*.
- 13 Miyake and Friedman (2012) see cognitive flexibility ('shifting') as involving shifting-specific EF and general EF (= cognitive inhibition, shared by all EFs) – see this chapter's Figure 4.5. They claim, however, that shifting-specific EF shows some evidence of *negative* associations with certain developmental indicators of cognitive

inhibition (which latter they see as a general factor partly underpinning EFs) – better inhibitory control, to some qualified extent, predicts poorer shifting-specific EF. On reflection, that’s kind of intuitive, isn’t it? Within the individual differences literature, Carson et al. (2003) established that decreased latent inhibition (the ability to ‘gate off’, attentionally, task-irrelevant stimuli) is associated with *increased* creative achievement (though in high-functioning individuals, not generally – where it may predict psychopathology). Decreased latent inhibition is also known to predict the exploratory cognitive virtues more generally, such as Big 5 O (openness to experience) – in Carson et al., *ibid*.

- 14 Whether this *token* would be an instance of a *type* – that is, whether EF selection is a sufficiently disjunctive, holistic product of an open system as to be not really (in the sense intended, in a sufficiently meaningful sense) *algorithmic* (fixed output for fixed inputs thereby mapping a finitely writeable, time transfer-invariant function) is of course a very well-worked debate in classic (functionalist) philosophy of mind. Whatever problems are held to attend functionalism on other grounds and in the context of other philosophical questions, I see no reason to retrench from that position’s standard emergentist stance in this domain.
- 15 In a sense, we do just this. We are the only organism with language and the only organism with (in any meaningful sense) culture (and technology). We are also, as noted, the only organism with anything like our level of autozoetic consciousness: our sense of self and self-relevant (counterfactual) model building (planning) across time. The socio-cultural extensions of frontal lobe function are extraordinary. Just as the extended mind hypothesis extends mind into the physical world, so the neo-Vygotskian tradition extends mind into culture, education and society: ‘One startling observation Premack made was that chimps do not engage in pedagogy. The flip side of the observation is that humans are the only primates that teach their young’ (Gazzaniga 2013: 12). Clark and Wilson note of Sue Savage-Rumbaugh’s bonobo, Kanzi, raised via a humanly imposed pedagogy using a symbolic keyboard: ‘The system that Kanzi plus his keyboard constitutes forms a cognitive system with memory and other cognitive capacities that seem qualitatively distinctive from that of other, unaugmented bonobos, capacities that are somewhere between those of humans and other apes. It is not simply that Kanzi’s enriched learning environment has restructured his neural wiring (although it has almost certainly done that too), but that his cognitive restructuring has proceeded through a potent cognitive extension involving these stable symbolic structures in his environment’ (Wilson and Clark 2009: 67).
- 16 ‘Motor structures’ for Fuster, after Koechlin, range from the most basic aspects of muscle movement mediated by the motor cortex, to the highest, most abstract, pure, agential cognition (e.g. Heisenberg cases) mediated significantly by the most anterior areas of the PFC (not intrinsically involving actual motor movement at all).

These structures are seen as ranging hierarchically, from those involving the most basic limb movement to those involving the most sophisticated levels of abstract goal-directed thought, but in every case involving multiple self-organizing, open-system feedback cycles from the perceptual to the motor and back.

- 17 Of course, this doesn't vitiate the importance of the PFC in EF, or cognitive agency per se. A strong statement of the role of the PFC as 'capital C' Controller is the *frontal executive hypothesis* – that all executive control processes are subserved by the frontal lobes (this is now known to be too strong). Less-strong would be *parieto-frontal integration theory* (P-FIT) – that attentional and executive processes are shared in a highly integrated fashion between frontal and parietal lobes. Probably matters are more complex than that, too. Nevertheless, even approaches conceding more to the holistic and emergentist traditions in neuropsychology acknowledge the critical role for the frontal lobes in the highest, most specifically human, rational, executive, metapsychological, reflexive, insightful, flexible and controlled thought. 'My PFC is not my "center of free will" but it is the neural broker of the highest transactions of myself with my environment, internal as well as external' (Fuster 2013:10).
- 18 The others are (3) attention regulation, (4) cognitive change/reappraisal and (5) response modification/modulation/suppression.
- 19 *Spacing*: the extraordinarily powerful 'spacing effect' – distributing practice a little and often and not 'massing'. *Interleaving* – introducing 'contextual interference' (e.g. practicing different swimming strokes (and different components of swimming strokes) in the same training session and not 'blocking'). *Generating* (e.g. generating answers, self-testing, elaborating on the material).
- 20 This is a slightly under-qualified statement. It may be read as rather too close to Giacomo, Rizzolatti et al.'s committed *Premotor Theory*, the truth of which is still in the balance (cf. Wu 2014). As regards attention, Wu (2014: 66, emphases in original) usefully distinguishes 'three possible formulations of the Premotor Theory: the *identity*, *causal* and *anatomic* formulations. The strongest thesis is the identity formulation: the neural circuits for preparing eye movement to location L just are the neural circuits for visual spatial attention to L. The causal formulation is weaker in that it endorses only a causally sufficient condition: normal preparatory processing for eye movement to location L is sufficient to cause visual spatial attention to L. Finally, the *anatomic* formulation holds that a brain region contains circuitry for preparatory eye movement activity and for spatial attention.' Suitably qualified, however, it is widely accepted that (in some sense) both online and offline efferent circuitry in the frontal eye fields is in intricate connection with (other) frontal lobe functions, in which role they mediate *movement of thought*.
- 21 Piaget (1952) contended that, for him, thought was impossible without writing. Wittgenstein absurdly, but rather splendidly, over-stated the Piaget–Feynman thesis



- thus: ‘this activity [thinking] is performed by the hand, when we think by writing, by the mouth and larynx, when we think by speaking (Wittgenstein 1958: 6).
- 22 To see how hard the notion of mental action may be pushed, certain brain scientists now take talk of ‘mental agility’ *literally*, this being a variable mediated by structures (especially the cerebellum) that mediate motor agility. Damage these structures and *dysmetria* results, but this may be a motor dysmetria (ataxia) or *dysmetria of thought* – an impairment of *mental agility* (e.g. Schmahmann 2004).
- 23 Notably, Johnson-Laird agrees with Evans here. Evans assimilates this (WM) to issues of control and volition. In seeking to understand the Type 1/Type 2 distinction ‘[w]e should not rely on the distinction between conscious and unconscious processing, as both old and new minds have aspects that are conscious and unconscious. ... The key difference is that the old mind operates through automated Type 1 systems that have mandatory outputs, whereas Type 2 systems are in some sense *volitional*: the new mind is capable of forming plans and carrying out intentions under controlled attention’ (Evans 2014: 133, emphasis in the original).
- 24 Note of their ‘bias’ item in the left-hand column: they are referring to ‘bias’ in the Kahneman–Tversky (heuristics and biases) sense: precisely not our sense – cf. next section. Keren and Schul (2009) in their influential polemic against two-system approaches note (correctly) the heterogeneous nature of such dichotomous lists as that offered by Evans and Stanovich (2013); adding to that heterogeneity a tendency among some theorists to associate System 1 cognition with hot cognition versus System 2 with cold. This is not found in the table I have taken from Evans and Stanovich (2013) and would seem to myself to be quite indefensible on its own terms. We have ‘old mind’ subcortical surges of affect, granted (Evans and Stanovich’s ‘basic emotions’); but we most certainly have ‘new mind’ great systems of ethics, normative epistemology and value theory more generally (Evans and Stanovich’s ‘complex emotions’): we have rigorous socioaffective cognition. Our executive functioning both mediates and is mediated by massive cultural scaffolding and unquestionably involves moral *thought*.
- 25 This identification (plus a non-essentialism about the binary distinction) allows us to accommodate an influential criticism from Keren and Schul (2009): that many of the processes subsumed under the two-process distinction are continuous, perhaps dimensional, rather than dichotomous or discrete. We have seen already how executive functioning segues into non-executive functioning – as, say, processes that once were mediated by the most anterior areas of the PFC become progressively more posterior and eventually perhaps no longer mediated by the PFC: via automaticity, proceduralization, etc. I am interested in, but not committed to, two-process/system/type theory; but were I to commit to it, I would not be an essentialist about it, and would want to defend it as a heuristically useful

- dimensional tendency to be kept in mind through psychological theorizing. Considered thus, some of Keren and Schul's criticisms are against a conception of the distinction that is excessively strong, abrupt and dichotomous. Note that the distinction between executive and non-executive functioning is clearly useful and well-founded whilst not being 'ontologizable' in any simply dichotomous way (e.g. the frontal lobe hypothesis is clearly too strong). Functionally, the distinction between executive and non-executive functioning requires the sort of explication we have seen over this chapter, not a 'definition' in any sense: it is a well-founded distinction for all that.
- 26 'Bias' is also used *passim* by researchers into 'hot' cognition and moral thought (e.g. Anderson et al. 1999).
- 27 Wu (2014) calls this phenomenon as it manifests in attention the Many-Many Problem.
- 28 Importantly, in light of (e.g. regress) issues shortly to be discussed, one may also *bias* one's *biases* – cf. Burgess et al.'s (2006) 'gateway hypothesis'.
- 29 I suspect that a strong reductionism, narrowly and locally and ontologically conceived, may be needed as a premise to make this argument work: 'I conceive that nothing taketh beginning from itself but from the action of some other immediate *agent* without itself' (Hobbes 1962: 271).
- 30 There are clear parallels with the classic compatibilist literature here. For three hundred years after Hobbes and prior to Frankfurt, one classic compatibilist after another denied the obvious descriptive truth that there were higher-order cognitions and conations – because, in the grip of a theory and seeing a regress beckoning, they felt that doing so would be insufficiently 'ultimate' for their purposes: 'I acknowledge this *liberty*, that I *can* do if I *will*; but to say, I can *will* if I *will*, I take to be an absurd speech' (Hobbes 1962: 246). This die-in-a-ditch tendency simply melted away when Frankfurt placed other resources at the compatibilists disposal – to the point where classic compatibilism is now (almost?) a museum piece. Yet the armchair psychological truths that Frankfurt adverted to (as to there being higher-order intentional states) should have been quite as obvious, descriptively, three hundred and fifty years ago as they are today.

## Chapter 5

- 1 A point originally made of happiness in the ethical literature (Sidgwick 1907: 405–6).
- 2 Of course, not if, by stipulation of the thought experiment, the demon is held able to manipulate the envatted agent's active thought processes – his decision-

## References

- Alexander, S. (1920) *Space, Time and Deity* (2 vols.), London: Macmillan.
- Allport, G. W. (1937) *Personality: A Psychological Interpretation*, New York: Holt, Rinehart, & Winston.
- Alston, W. (1985) 'Concepts of epistemic justification,' *Monist*, 68(1): 57–89, reprinted in Alston, 1989b, 81–115.
- Alston, W. (1988) 'The deontological conception of epistemic justification,' in J. E. Tomberlin (ed), *Philosophical Perspectives: Vol. 2, Epistemology*, Atascadero: Ridgeview, 257–99. Reprinted in Alston, 1989b, 115–53.
- Alston, W. (1989a) 'Internalism and externalism in epistemology,' in W. Alston, *Epistemic Justification*, Ithaca: Cornell University Press, 185–226.
- Alston, W. (1989b) *Epistemic Justification*, Ithaca: Cornell University Press.
- Anderson, A. (1957) 'The formal analysis of normative concepts,' *American Sociological Review*, 22: 9–17.
- Anderson, S., A. Bechara, H. Damasio, D. Tranel and A. Damasio, (1999) 'Impairment of social and moral behaviour related to early damage in human prefrontal cortex,' *Nature Neuroscience*, 2: 1032–37.
- Anderson, V., E. Northam, J. Hendy and J. Wrenall (2001) *Developmental Neuropsychology: A Clinical Approach*, New York: Psychology Press.
- Annas, J. (2003) 'The structure of virtue,' in L. Zagzebski and M. de Paul (eds), *Intellectual Virtue: Perspectives from Ethics and Epistemology*, Oxford: Oxford University Press, 15–33.
- Anscombe, G. E. M. (1971) *Causality and Determination: An Inaugural Lecture*, Cambridge: Cambridge University Press; reprinted in G. E. M. Anscombe (1981) *Metaphysics and the Philosophy of Mind (The Collected Philosophical Papers of G. E. M. Anscombe)* (vol. 2), Minneapolis: University of Minnesota Press, 133–47.
- Anscombe, G. E. M. (1981) 'A reply to Mr. C.S. Lewis's argument that 'naturalism' is self-refuting,' in *Metaphysics and the Philosophy of Mind (The Collected Philosophical Papers of G. E. M. Anscombe)*, (vol. 2), Minneapolis: University of Minnesota Press, 224–32.
- Aristotle (1941) *Metaphysics*, in Richard McKeon (ed), *The Basic Works of Aristotle*; tr. W. D. Ross, 1928, 2nd edn. New York: Random House.
- Aristotle (1999) *Physics Book VIII*, Clarendon Aristotle Series, tr. D. W. Graham, Oxford: Clarendon Press.
- Aristotle (2000) *Nicomachean Ethics*, tr. R. Crisp, Cambridge: Cambridge University Press.

- Atkinson, R. C. and R. M. Shiffrin (1968) 'Human memory: A proposed system and its control processes', in K. W. Spence and J. T. Spence (eds), *The Psychology of Learning and Motivation* (vol. 2). New York: Academic Press, 89–195.
- Audi, R. (2001) 'Doxastic voluntarism and the ethics of belief', in M. Steup (ed), *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility and Virtue*, Oxford: Oxford University Press, 95–111.
- Baddeley, A. and G. Hitch (1974) 'Working memory', in G. H. Bower (ed), *The Psychology of Learning and Motivation: Advances in Research and Theory* (vol. 8), New York: Academic Press, 47–89.
- Badre, D. and M. D'Esposito (2009) 'Is the rostro-caudal axis of the frontal lobe hierarchical?' *National Review of Neuroscience*, 10: 659–69.
- Balaguer, M. (2010) *Free Will as an Open Scientific Problem*, Cambridge, MA: MIT Press.
- Bales, R. E. (1971) 'Act-Utilitarianism: Account of right-making characteristics or decision-making procedure?' *American Philosophical Quarterly*, 8: 257–65.
- Ballard, D., M. Hayhoe, P. Pook and R. Rao (1997) 'Deictic codes for the embodiment of cognition', *Behavioral and Brain Sciences*, 20: 723–67.
- Banich, M. (2004) *Cognitive Neuroscience and Neuropsychology*, Boston: Houghton Mifflin.
- Bannister, D. and F. Fransella (1971) *Inquiring Man: The Theory Of Personal Constructs*, Harmondsworth: Penguin.
- Barkley, R. (1997) 'Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD', *Psychological Bulletin*, 121(1): 65–94.
- Barkley, R. (2012) *Executive Functions*, New York: Guilford Press.
- Bartlett, F. C. (1932) *Remembering*, Cambridge: Cambridge University Press.
- Bealer, G. (1996a) 'A priori knowledge and the scope of philosophy', *Philosophical Studies*, 81: 121–42.
- Bealer, G. (1996b) 'On the possibility of philosophical knowledge', *Noûs*, 30 (Supplement): *Philosophical Perspectives*, 10, Metaphysics 1–34.
- Bechara, A. and A. Damasio (2005) 'The somatic marker hypothesis: A neural theory of economic decision', *Games and Economic Behavior*, 52: 336–72.
- Bechara, A., A. Damasio, H. Damasio and S. Anderson (1994) 'Insensitivity to future consequences following damage to human prefrontal cortex', *Cognition*, 50(1–3): 7–15.
- Bechara, A., H. Damasio, D. Tranel and A. Damasio (2005) 'The Iowa gambling task and the somatic marker hypothesis: Some questions and answers', *Trends in Cognitive Sciences*, 9(4): 159–62.
- Beebe, H. (2001) 'Transfer of warrant, begging the question and semantic externalism', *Philosophical Quarterly*, 51: 356–74.
- Bergmann, M. (2006) *Justification without Awareness*, Oxford: Clarendon Press.
- Berofsky, B. (2002) 'Ifs, cans and free will: The issues', in R. Kane (ed), *Oxford Handbook of Free Will*, Oxford: Oxford University Press, 181–201.

- Best, J., P. Miller and L. Jones (2009) 'Executive functions after age 5: Changes and correlates', *Developmental Review*, 29(3): 180–200.
- Bjork, R. (1994) 'Memory and metamemory considerations in the training of human beings', in J. Metcalfe and A. Shimamura (eds), *Metacognition: Knowing about Knowing*. Cambridge, MA: MIT Press, 185–205.
- Bjork, R., J. Dunlosky and N. Kornell (2013) 'Self-regulated learning: Beliefs, techniques, and illusions', *Annual Review of Psychology*, 64: 417–44.
- Bobzien, S. (2001) *Determinism and Freedom in Stoic Philosophy*, Oxford: Oxford University Press.
- BonJour, L. (1985) *The Structure of Empirical Knowledge*, Cambridge, MA: Harvard University Press.
- BonJour, L. (1998) *In Defense of Pure Reason*, Cambridge: Cambridge University Press.
- BonJour, L. (2003) *Epistemic Justification: Internalism vs. Externalism, Foundations vs Virtues*, Oxford: Blackwell.
- BonJour, L. (2010) 'Recent work on the internalism-externalism controversy', in J. Dancy, E. Sosa and M. Steup (eds), *A Companion to Epistemology*, 2nd edn. Oxford: Blackwell, 33–42.
- Borkowsky, J. and J. Burke (1996) 'Theories, models and measurements of executive functioning: An information processing perspective', in G. R. Lyon and N. A. Krasnegor (eds), *Attention, Memory and Executive Function*, Baltimore: Paul H. Brookes, 235–61.
- Brink, D. (1986) 'Utilitarian morality and the personal point of view', *Journal of Philosophy*, 83(8): 417–38.
- Broad, C. D. (1925) *The Mind and Its Place in Nature*, London: Routledge & Kegan Paul.
- Broad, C. D. (1952) 'Determinism, indeterminism and libertarianism', in *Ethics and the History of Philosophy, Selected Essays*, London: Routledge & Kegan Paul, 195–217.
- Burgess, P., J. Simons, I. Dumontheil and S. Gilbert (2006) 'The gateway hypothesis of rostral prefrontal cortex (area 10) function', in J. Duncan, L. Phillips and P. McLeod (eds), *Measuring the Mind: Speed, Control & Age*, Oxford: Oxford University Press, 217–48.
- Burnyeat, M. (1976) 'Protagoras and self-refutation in later Greek philosophy', *Philosophical Review*, 85(1): 44–69.
- Campbell, C. A. (1957) 'Has the self 'free will'?' in *On Selfhood and Godhood*, London: Allen & Unwin, 158–78.
- Carson, S., J. Peterson and D. Higgins (2003) 'Decreased latent inhibition is associated with increased creative achievement in high-functioning individuals', *Journal of Personality & Social Psychology*, 85(3): 499–506.
- Chisholm, R. (1956a) 'Epistemic statements and the ethics of belief', *Philosophy and Phenomenological Research*, XVI: 4.
- Chisholm, R. (1956b) 'The concept of empirical evidence', *Journal of Philosophy*, 53.
- Chisholm, R. (1957) *Perceiving: A Philosophical Study*, Ithaca: Cornell.

- Chisholm, R. (1986a) 'Self profile', in R. J. Bogdan (ed), *Roderick M. Chisholm*, Dordrecht: D. Reidel, 3–77.
- Chisholm, R. (1986b) 'The place of epistemic justification', *Philosophical Topics*, 14: 85–92.
- Chisholm, R. (1988) 'The indispensability of internal justification', *Synthese*, 74: 285–96.
- Chisholm, R. (1989) *Theory of Knowledge*, 3rd edn, Englewood Cliffs: Prentice Hall.
- Chomsky, N. (1988) *Language and Problems of Knowledge*, Cambridge, MA: MIT Press.
- Chomsky, N. (1995) 'Language and nature', *Mind*, 104, January.
- Chomsky, N. (2004) 'Things no amount of learning can teach', reprinted in C. Otero (ed), *Language and Politics*, 2nd edn, Edinburgh: A K Press; originally *Omni*, 6:11, November 1983.
- Chomsky, N. (2016) *What Kind of Creatures are We?* New York: Columbia University Press.
- Chrisman, M. (2008) 'Ought to believe', *Journal of Philosophy*, 105(7): 346–70.
- Churchland, P. S. (1981) 'Is determinism self-refuting?', *Mind*, 90: 99–101.
- Cicero (1942) *de Fato*, tr. H. Rackham, Heinemann Ltd. Cambridge, MA: Harvard University Press.
- Clark, A. (2008) *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*, Oxford: Oxford University Press.
- Clark, A. (2012) 'Embodied, embedded, and extended cognition', in K. Frankish and W. M. Ramsey (eds), *The Cambridge Handbook of Cognitive Science*, Cambridge: Cambridge University Press, 275–91.
- Clarke, M. (1986) 'Doxastic voluntarism and forced belief', *Philosophical Studies*, 50(1): 39–51.
- Clarke, R. (1995) 'Toward a credible agent-causal account of free will', in T. O'Connor (ed), *Agents, Causes, and Events: Essays on Indeterminism and Free Will*, New York: Oxford University Press, 201–15.
- Clement, C. and R. Falmange (1986) 'Logical reasoning, world knowledge, and mental imagery: Interconnections in cognitive processes', *Memory & Cognition*, 14(4): 299–307.
- Clifford, W. K. (1999) 'The ethics of belief', in T. Madigan (ed), *The Ethics of Belief and Other Essays*, Amherst and New York: Prometheus Books.
- Cohen, L. J. (1981) 'Can human irrationality be experimentally demonstrated?', *Behavioral and Brain Sciences*, 4: 317–70.
- Cole, M. and S. Scribner (1974) *Culture and Thought; A Psychological Introduction*, Wiley.
- Compton, A. (1939) *The Freedom of Man*, 3rd edn, New Haven: Yale University Press.
- Conee, E. and R. Feldman (2001) 'Internalism defended', in H. Kornblith (ed), *Internalism and Externalism*, Oxford: Blackwell, 231–60.
- Conway, A., S. Getz, B. Macnamara and P. Engel de Abreu (2011) 'Working memory & intelligence', in R. Sternberg and S. Kaufman (eds), *The Cambridge Handbook of Intelligence*, Cambridge: Cambridge University Press, 394–418.

- Cooper, L. and R. Shepard (1973) 'The time required to prepare for a rotated stimulus', *Memory & Cognition*, 1: 246–50.
- Copp, D. (2003) "Ought" implies "can", blameworthiness, and alternative possibilities', in D. Widerker and M. McKenna (eds), *Freedom, Responsibility and Agency: Essays on the Importance of Alternative Possibilities*, Aldershot: Ashgate Press, 265–99.
- Copp, D. (2008) "Ought" implies "can" and the derivation of the principle of alternative possibilities', *Analysis*, 68(1): 67–75.
- Danto, A. (1987) 'Approaching the end of art', in *The State of the Art*, New York: Prentice Hall Press, 202–18.
- Davidson, D. (1980) 'Mental events', in D. Davidson (ed), *Essays on Actions and Events*, Oxford: Clarendon Press, 207–24.
- Davies, M. (2004) 'Epistemic entitlement, warrant transmission and easy knowledge', *Proceedings of the Aristotelian Society Supplementary*, 78: 213–45.
- Delis, D., E. Kaplan and N. Kramer (2001) *Delis–Kaplan Executive Function System*, Odessa: Psychological Assessment Resources.
- Dennett, D. (1981) 'Making sense of ourselves', *Philosophical Topics*, 12: 63–81.
- Dennett, D. (1984) *Elbow Room: the Varieties of Free Will Worth Wanting*, Oxford: Clarendon Press.
- Descartes, R. (1931) *Philosophical Works of Descartes Vol. 1*, E. Haldane and G. Ross (eds), Cambridge: Cambridge University Press.
- Devitt, M. and G. Rey (1991) 'Transcending transcendentalism', *Pacific Philosophical Quarterly*, 72: 87–100.
- Diamond, A. (2013) 'Executive functions', *Annual Review of Psychology*, 64: 135–68.
- Diogenes, L. (1925) *The Lives and Opinions of Eminent Philosophers* (vol. II), tr. R. D. Hicks, Loeb Classical Library.
- Duncker, K. (1945) 'On problem-solving', tr. S. Lynne, *Psychological Monographs*, 58(5): i–113.
- Dunlosky, J., M. J. Serra and J. M. C. Baker (2007) 'Metamemory', in F. T. Durso, R. S. Nickerson, S. T. Dumais, S. Lewandowsky and T. J. Perfect (eds), *Handbook of Applied Cognition*, 2nd edn, Chichester: John Wiley & Sons Ltd, 137–60.
- Edelman, G. and J. Gally (2001) 'Degeneracy and complexity in biological systems', *Proceedings of the National Academy of Sciences*, 98: 13763–68.
- Edwards, J. (2007) *Freedom of the Will*, [1768] New York: Cosimo Classics.
- Elliot, R. (2003) 'Executive functions and their disorders', *British Medical Bulletin*, 65: 49–59.
- Elqayam, S. (2011) 'Grounded rationality: A relativist framework for normative rationality', in: K. Manktelow, D. Over and S. Elqayam (eds), *The Science of Reason: A Festschrift in Honour of Jonathan St.B.T. Evans*, Hove: Psychology Press, 397–420.
- Elqayam, S. (2012) 'Grounded rationality: Descriptivism in epistemic context', *Synthese*, 189: 39–49.
- Epicurus (1964) *Letters, Principle Doctrines and Vatican Sayings*, tr. R. M. Geer, Indianapolis and New York: Bobbs-Merrill.

- Erasmus (2005) in Erasmus & Luther (2005) *Discourse on Free Will*, tr. & ed. E. F. Winter, Continuum.
- Ericsson, K., R. Krampe and C. Tesch-Romer (1993) 'The role of deliberate practice in the acquisition of expert performance', *Psychological Review*, 100(3): 361–406.
- Evans, J. St. B. T. (2014): 'Two minds rationality', *Thinking & Reasoning*, 20(2): 129–46.
- Evans, J. St. B. T. and K. E. Stanovich (2013) 'Dual-process theories of higher cognition: Advancing the debate', *Perspectives in Psychological Science*, 8: 223–41.
- Ewing, A. C. (1937) 'Meaninglessness', *Mind*, 46(6): 347–64.
- Feldman, R. (1985) 'Reliability and justification', *Monist*, 68: 150–74.
- Feldman, R. (2001) 'Voluntary belief and epistemic evaluation', in M. Steup (ed), *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility and Virtue*, Oxford: Oxford University Press, 77–92.
- Fischer, J. M. and M. Ravizza (1998) *Responsibility and Control: A Theory of Moral Responsibility*, Cambridge: Cambridge University Press.
- Fischer, J. M. and M. Ravizza (2000) 'Précis of responsibility and control: A theory of moral responsibility', *Philosophy & Phenomenological Research*, 61(2): 441–6.
- Fischer, J. M. (1994) *The Metaphysics of Free Will*, Oxford: Blackwell.
- Fischer, J. M. (1999) 'Recent work on moral responsibility', *Ethics*, 110: 93–139.
- Fischer, J. M. (2006a) 'Free will and moral responsibility' in D. Copp (ed), *The Oxford Handbook of Ethical Theory*, Oxford: Oxford University Press, 321–54.
- Fischer, J. M. (2006b) *My Way: Essays on Moral Responsibility*, Oxford: Oxford University Press.
- Fleming, K., S. Heintzelman and B. Bartholow (2016) 'Specifying associations between conscientiousness and executive functioning: Mental set shifting, not prepotent response inhibition or working memory updating', *Journal of Personality*, 84: 348–60.
- Fodor, J. (1974) 'Special sciences, or the disunity of science as a working hypothesis', *Synthese*, 28: 77–115.
- Fodor, J. (1988) *Psychosemantics*, Cambridge, MA: MIT Press.
- Foley, R. (1993) *Working without a Net: A Study of Ego-centric Epistemology*, New York and Oxford: Oxford University Press.
- Foley, R. (2001) *Intellectual Trust in Oneself and Others*, New York: Cambridge University Press.
- Foley, R. (2004) 'A trial separation between the theory of knowledge and the theory of justified belief', in J. Greco (ed), *Ernest Sosa and His Critics*, Oxford: Wiley, 59–72.
- Foot, P. (1978) 'Morality as a system of hypothetical imperatives', in *Virtues and Vices*, Oxford: Blackwell, 157–73.
- Frankfurt, H. (1969), 'Alternate possibilities and moral responsibility', *The Journal of Philosophy*, 66: 829–39; in his *The Importance of What We Care about*, Cambridge: Cambridge University Press (1988).
- Frankfurt, H. (1971) 'Freedom of the will and the concept of a person', *The Journal of Philosophy*, 68; in his *The Importance of What We Care about*, Cambridge: Cambridge University Press (1988).



- Frankfurt, H. (1975) 'Three concepts of free action' *Proceedings of the Aristotelian Society Supplementary*, 44: 113–25; in his *The Importance of What We Care About*, Cambridge: Cambridge University Press (1988).
- Frankfurt, H. (1983) 'What we are morally responsible for', in *The Importance of What We Care about*, Cambridge: Cambridge University Press, 95–103.
- Frankfurt, H. (1994) 'An alleged asymmetry between actions and omissions', *Ethics*, 104: 620–3.
- Frankfurt, H. (2003) 'Some thoughts concerning PAP', in D. Widerker and M. McKenna (eds), *Freedom, Responsibility and Agency: Essays on the Importance of Alternative Possibilities*, Aldershot: Ashgate Press, 339–45.
- Frankish, K. (2011) 'Conscious thinking, acceptance, and self-deception', *Behavioral and Brain Sciences*, 34(1): 20–1.
- Frege, G. (1980) *Foundations of Arithmetic*, tr. J. L. Austin, Oxford: Blackwell.
- Friedman, N., A. Miyake, S. Young, J. DeFries, R. Corley and J. Hewitt (2008) 'Individual differences in executive functions are almost entirely genetic in origin', *Journal of Experimental Psychology: General*, 137(2): 201–25.
- Fuster, J. (2004) 'Upper processing stages of the perception–action cycle', *Trends in Cognitive Sciences*, 8(4): 143–5.
- Fuster, J. (2013) *The Neuroscience of Freedom and Creativity*, Cambridge: Cambridge University Press.
- Gazzaniga, M. (2013) 'Shifting gears: Seeking new approaches for mind/brain mechanisms', *Annual Review of Psychology*, 64: 1–20.
- Gellius, A. (1927) *Attic Nights of Aulus Gellius*, tr. J. C. Rolfe, Cambridge, MA: Harvard University Press.
- Gemes, K. (2006) 'Nietzsche on free will', *Proceedings of the Aristotelian Society*, 80: 321–38.
- Gick, M. and K. Holyoak (1980) 'Analogical problem solving', *Cognitive Psychology*, 12: 306–55.
- Gick, M. and K. Holyoak (1983) 'Schema induction and analogical transfer', *Cognitive Psychology*, 15: 1–28.
- Ginet, C. (1996) 'In defense of the principle of alternative possibilities', *Philosophical Perspectives*, 10: 403–17.
- Ginet, C. (2001) 'Deciding to believe', in M. Steup (ed), *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility and Virtue*, Oxford: Oxford University Press, 63–76.
- Ginet, C. (2002a) 'Freedom, responsibility and agency', in R. Kane (ed), *Free Will*, Blackwell Readings in Philosophy, Oxford: Blackwell, 206–21.
- Ginet, C. (2002b) 'Reasons explanation of action: Causalist versus noncausalist accounts', in R. Kane (ed), *The Oxford Handbook of Free Will*, Oxford: Oxford University Press, 386–405.
- Goldman, A. (1967) 'A causal theory of knowing', *Journal of Philosophy*, 64(12): 357–72.

- Goldman, A. (1969) 'The compatibility of mechanism and purpose', *Philosophical Review*, 78: 468–82.
- Goldman, A. (1980) 'The internalist conception of justification', *Midwest Studies in Philosophy* 5: 27–52.
- Goldman, A. (1986) *Epistemology and Cognition*, Cambridge, MA: Harvard University Press.
- Goldman, A. (1988) 'Strong and weak justification', in J. E. Tomberlin (ed), *Philosophical Perspectives* 2, *Epistemology*, 51–69.
- Goldman, A. (2001) 'The unity of the epistemic virtues' in A. Fairweather and L. Zagzebski (eds), *Virtue Epistemology: Essays On Epistemic Virtue And Responsibility*, Oxford: Oxford University Press, 30–48.
- Goldman, A. (2007) 'Philosophical intuitions: Their target, their source, and their epistemic status', *Grazer Philosophische Studien*, 74: 1–26
- Goldman, A. (2009) 'Internalism, externalism and the architecture of justification', *Journal of Philosophy*, 106(6): 1–30.
- Grant, E. and M. Spivey (2003) 'Eye movements and problem solving: Guiding attention guides thought', *Psychological Science*, 14: 462–6.
- Greco, J. (2002) 'Virtues in epistemology', in P. Moser (ed), *The Oxford Handbook of Epistemology*, Oxford: Oxford University Press, 287–315.
- Greco, J. (2005) 'Justification is not internal', in M. Steup and E. Sosa (eds), *Contemporary Debates in Epistemology*, Oxford: Blackwell, 257–70.
- Greene, J. (2007) 'Why are VMPFC patients more utilitarian?: A dual-process theory of moral judgment explains', *Trends in Cognitive Sciences*, 11(8): 322–3.
- Greene, J., L. Nystrom, A. Engell, J. Darley and J. Cohen (2004) 'The neural bases of cognitive conflict and control in moral judgment', *Neuron*, 44, 389–400.
- Greene, J., R. Sommerville, L. Nystrom, J. Darley and J. Cohen (2001) 'An fMRI investigation of emotional engagement in moral judgment', *Science*, 293: 2105–8.
- Gross, J. and R. Thompson (2007) 'Emotion regulation: Conceptual foundations', in J. J. Gross (ed), *Handbook of Emotion Regulation*, New York: Guilford Press, 3–24.
- Grünbaum, A. (1971) 'Free will and laws of human behaviour', *American Philosophical Quarterly*, 8(4): 299–317.
- Haier, R. (2017) *The Neuroscience of Intelligence*, New York: Cambridge University Press.
- Haji, I. (1999) 'Moral anchors and control', *Canadian Journal of Philosophy*, 29(2): 175–203.
- Haji, I. (2008) 'Reflections on the incompatibilist's direct argument', *Erkenntnis*, 68: 1–19.
- Hasker, W. (1973) 'The transcendental refutation of determinism', *Southern Journal of Philosophy*, 11: 175–83.
- Heil, J. (1983) 'Doxastic agency', *Philosophical Studies*, 43: 355–64.
- Heisenberg, W., cited in T. Powers (1993) *Heisenberg's War: The Secret History of the German Bomb*, New York: Knopf Press.

- Hieronymi, P. (2005) 'The wrong kind of reason', *Journal of Philosophy*, 102(9): 437–57.
- Hieronymi, P. (2008) 'Responsibility for believing', *Synthese*, 161(3): 357–73.
- Hobbes, T. (1962) 'Of liberty and necessity, Lord Newcastle's opinions examined', in R. S. Peters (ed), *Body, Man, & Citizen: Selections from Hobbes*, New York: Collier Books, 245–60.
- Hobson, P. and L. Leeds (2001) 'Executive functioning in older people', *Reviews in Clinical Gerontology*, 11: 361–72.
- Hofmann, W., B. Schmeichel and A. Baddeley (2012) 'Executive functions and self-regulation', *Trends in Cognitive Sciences*, 16(3): 174–80.
- Hue, A. and E. Southby-Taylor (2005) *The Next Moon*, Harmondsworth: Penguin.
- Hunt, E. and S. Shabo (2013) 'Frankfurt cases and the (in)significance of timing: A defense of the buffering strategy', *Philosophical Studies*, 164: 599–622.
- James, W. (1897a) 'The will to believe', in *The Will to Believe and Other Essays in Popular Philosophy*, London: Longmans, Green & Co.; reprinted Dover Publications, New York, 1956, 1–31.
- James, W. (1897b) 'The dilemma of determinism', in *The Will to Believe and Other Essays in Popular Philosophy*, London: Longmans, Green & Co.; reprinted Dover Publications, New York, 1956, 145–83.
- Johnson-Laird, P. and M. Steedman (1978) 'The psychology of syllogisms', *Cognitive Psychology*, 10, 64–99.
- Johnson-Laird, P. (2001) 'Mental models, deductive reasoning, and the brain', in M. S. Gazzaniga (ed), *The Cognitive Neurosciences*, Cambridge: MIT Press, 999–1008.
- Jordan, J. (1969) 'Determinism's dilemma', *Review of Metaphysics*, 23: 48–66.
- Jurado, M. and M. Rosselli (2007) 'The elusive nature of executive functions: A review of our current understanding', *Neuropsychological Review*, 17(3): 213–33.
- Kane, R. (1996) *The Significance of Free Will*, Oxford: Oxford University Press.
- Kane, R. (2000) 'The dual regress of free will and the role of alternative possibilities', *Philosophical Perspectives*, 14: 57–79.
- Kane, R. (2002) 'Free will: New directions for an ancient problem', in R. Kane (ed), *Free Will*, Blackwell Readings in Philosophy, Oxford: Blackwell, 222–48.
- Kant, I. (1933) *Critique of Pure Reason*, tr. N. Kemp-Smith, Basingstoke, Hants and London: Macmillan.
- Kant, I. (1964) *Groundwork of the Metaphysic of Morals*, tr. H. J. Paton, New York: Harper Torchbooks.
- Kant, I. (2004) *Critique of Practical Reason*, tr. Abbott, Whitefish: Kessinger.
- Kapitan, T. (2002) 'A master argument for incompatibilism?', in R. Kane (ed), *The Oxford Handbook of Free Will*, Oxford: Oxford University Press, 127–57.
- Keren, G. and Y. Schul (2009) 'Two is not always better than one: A critical evaluation of two-system theories', *Perspectives on Psychological Science*, 4: 533–50.
- Kim, J. (1989) 'Mechanism, purpose, and explanatory exclusion', *Philosophical Perspectives*, 3: 77–108.

- Kim, J. (1992) “Downward Causation” in emergentism and nonreductive physicalism, in A. Beckermann, H. Flohr and J. Kim (eds), *Emergence or Reduction?: Essays on the Prospects of Nonreductive Physicalism*, Walter de Gruyter & Co, 119–38.
- Kim, J. (2000) *Mind in a Physical World*, Bradford Books: MIT Press.
- Koechlin, E., C. Ody and F. Kouneiher (2003) ‘The architecture of cognitive control in the human prefrontal cortex’, *Science*, 302: 1181–5.
- Kornblith, H. (1983) ‘Justified belief and epistemically responsible action’, *Philosophical Review*, 92(1): 33–48.
- Kornblith, H. (2001) ‘Epistemic obligation and the possibility of internalism’, in A. Fairweather and L. Zagzebski (eds), *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*, Oxford: Oxford University Press, 231–48.
- Kornblith, H. (2007) ‘Naturalism and intuitions’, *Grazer Philosophische Studien*, 72: 27–49. Reprinted in C. Beyer and A. Burri (eds) (2007), *Philosophical Knowledge: Its possibility and Scope*, Amsterdam: Rodopi.
- Kornblith, H. (2012) *On Reflection*, Oxford: Oxford University Press.
- Kripke, S. (1980) *Naming and Necessity*, Cambridge, MA: Harvard University Press.
- Lafleche, G. and M. Albert (1995) ‘Executive function deficits in early Alzheimer’s disease’, *Neuropsychology*, 9: 313–20
- Leibniz, G. W. (2005) *Theodicy: Essays on the Goodness of God, the Freedom of Man and the Origin of Evil*, tr. E. M. Huggard, London: Routledge.
- Leite, A. (2004) ‘On justifying and being justified’, *Philosophical Issues, Epistemology*, 14: 219–53.
- Lewes, G. H. (1875) *Problems of Life and Mind* (First Series) (vol. 2), London: Kegan Paul, Trench, Trübner & Co.
- Lewis, C. S. (1947/1960) *Miracles*, 1st/2nd edn, London and Glasgow: Collins/Fontana.
- Lezak, M. (1983) *Neuropsychological Assessment*, 2nd edn, New York: Oxford University Press.
- Locke, J. (1975) *Essay Concerning Human Understanding*, P. H. Niddich (ed), Oxford: Oxford University Press.
- Lockie, R. (1998) ‘What’s wrong with moral internalism’, *Ratio*, 11(1): 14–36.
- Lockie, R. (2003a) ‘Depth psychology and self-deception’, *Philosophical Psychology*, 16(1): 127–48.
- Lockie, R. (2003b) ‘Relativism and reflexivity’, *International Journal of Philosophical Studies*, 11(3): 319–39.
- Lockie, R. (2003c) ‘Transcendental arguments against eliminativism’, *British Journal for the Philosophy of Science*, 54(4): 569–89.
- Lockie, R. (2004) ‘Knowledge, provenance and psychological explanation’, *Philosophy*, 79(3): 421–33.
- Lockie, R. (2006) ‘Response to Anders Tolland’s “iterated non-refutation: Robert Lockie on relativism”’, *International Journal of Philosophical Studies*, 14(2): 245–54.
- Lockie, R. (2008) ‘Problems for virtue theories in epistemology’, *Philosophical Studies*, 138(2): 169–91.

- Lockie, R. (2014a) 'Three recent Frankfurt cases', *Philosophia*, 42(4): 1005–32.
- Lockie, R. (2014b) 'The epistemology of neo-gettier epistemology', *South African Journal of Philosophy*, 33(2): 247–58.
- Lockie, R. (2014c) 'The regulative and the theoretical in epistemology', *Abstracta*, 8(1): 3–14.
- Lockie, R. (2016a) 'Perspectivism, deontologism and epistemic poverty', *Social Epistemology*, 30(2): 133–49.
- Lockie, R. (2016b) 'Response to Elqayam, Nottelmann, Peels and Vahid on my paper "perspectivism, deontologism and epistemic poverty"', *Social Epistemology Review and Reply Collective*, 5(3): 21–47.
- Lockie, R. (forthcoming) 'Virtue Epistemology and the Sources of Epistemic Value, in H. Battaly (ed), *Routledge Handbook of Virtue Epistemology*, Abingdon: Routledge.
- Lucas, J. R. (1970) *The Freedom of the Will*, Oxford: Oxford University Press.
- Lucretius (1947) *De Rerum Natura*, vol. III; tr. and commentary C. Bailey, Oxford: Oxford University Press.
- Luria, A. (1979) *The Making of Mind: A Personal Account of Soviet Psychology*, Cambridge, MA: Harvard University Press.
- Luria, A. (1973) *The Working Brain*, tr. B. Haigh, New York: Basic Books.
- Luria, A. (1976) *Cognitive Development, its Cultural and Social Foundations*. Cambridge, MA: Harvard University Press.
- Luther (1969) *De Servo Arbitrio*, tr. and ed. P. S. Watson in collaboration with B. Drewery, in Luther and Erasmus (1969), *Free Will and Salvation*, London: Westminster John Knox Press.
- MacDonald, K. (2008) 'Effortful control, explicit processing, and the regulation of human evolved predispositions', *Psychological Review*, 115(4): 1012–31.
- Maier, N. (1930) 'Reasoning in humans. I. On direction', *Journal of Comparative Psychology*, 10(2): 115–43.
- Maier, N. (1931) 'Reasoning in humans. II. The solution of a problem and its appearance in consciousness', *Journal of Comparative Psychology*, 12(2): 181–94.
- Malcolm, N. (1968) 'The conceivability of mechanism', *Philosophical Review*, 77(1): 45–72.
- Martin, W. (2009) 'Ought but cannot', *Proceedings of the Aristotelian Society*, 109(1): 103–28.
- McCall, S. and E. J. Lowe (2005) 'Indeterminist free will', *Philosophy & Phenomenological Research*, 70(3): 681–90.
- McGinn, C. (1993) *Problems in Philosophy: The Limits of Inquiry*, Oxford: Blackwell.
- McHugh, C. (2013) 'Epistemic responsibility & doxastic agency', *Philosophical Issues*, 23: 132–57.
- McKenna, M. (2008) 'Saying good-bye to the direct argument the right way', *Philosophical Review*, 117(3): 349–83.
- McLaughlin, B. P. (1992) 'The rise and fall of British emergentism', in A. Beckermann, H. Flohr and J. Kim (eds), *Emergence or Reduction?*, Berlin: Walter de Gruyter & Co, 49–93.

- Mele, A. (2006) *Free Will and Luck*, Oxford: Oxford University Press.
- Mill, J. S. (1843) *System of Logic*, London: Longmans, Green, Reader, and Dyer.
- Miller, E. and J. Cohen (2001) 'An integrative theory of prefrontal cortex function', *Annual Review of Neuroscience*, 24: 167–202.
- Miller, G. A. (1956) 'The magical number seven, plus or minus two: Some limits on our capacity for processing information', *Psychological Review*, 63(2): 81–97.
- Miller, G., E. Galanter and K. Pribram (1960) *Plans and the Structure of Behavior*, New York: Henry Holt and Company.
- Mischel, W., Y. Shoda and M. Rodriguez (1989) 'Delay of gratification in children', *Science*, 244(4907): 933–8.
- Miyake, A. and N. Friedman (2012) 'The nature and organization of individual differences in executive functions: Four general conclusions', *Current Directions in Psychological Science*, 21(1): 8–14.
- Moll, J. and R. de Oliveira-Souza (2007) 'Moral judgments, emotions, and the utilitarian brain', *Trends in Cognitive Sciences*, 11(8): 319–21.
- Monsell, S. (2003) 'Task switching', *Trends in Cognitive Sciences*, 7: 134–40.
- Moya, C. (2006) *Moral Responsibility*, Oxford: Routledge.
- Moya, C. (2011) 'On the very idea of a robust alternative', *Crítica, Revista Hispanoamericana Filosofía*, 41(August): 3–26.
- Nagel, E. (1961) *The Structure of Science: Problems in the Logic of Scientific Explanation*, London: Routledge and Kegan Paul.
- Nagel, E. (1963) 'Wholes, sums, and organic unities', in D. Lerner (ed), *Parts and Wholes*, New York: The Free Press, 135–55.
- Nagel, T. (1986) *The View from Nowhere*, Oxford: Oxford University Press.
- Nagel, T. (1997) *The Last Word*, Oxford: Oxford University Press.
- Neisser, U. (1967) *Cognitive Psychology*, Vancouver: Appleton, Century, Crofts.
- Nelkin, D. (2004a) 'Deliberative alternatives', *Philosophical Topics*, 32(1–2): 215–40.
- Nelkin, D. (2004b) 'The sense of freedom', in J. Campbell, M. O'Rourke and D. Shier (eds), *Freedom and Determinism*, Cambridge MA: MIT Press, 105–34.
- Nelkin, D. (2011) *Making Sense of Freedom and Responsibility*, Oxford: Oxford University Press.
- Nelson, T. and L. Narens (1990) 'Metamemory: A theoretical framework and new findings', *The Psychology of Learning and Motivation*, 26: 125–73.
- Nietzsche, F. (1973) *Beyond Good and Evil*, tr. R. J. Hollingdale, Harmondsworth: Penguin.
- Nietzsche, F. (1996) *Human, all too Human*, tr. R. J. Hollingdale, Cambridge: Cambridge University Press.
- Nisbett, R. E., J. Aronson, C. Blair, W. Dickens, J. Flynn, D. F. Halpern and E. Turkheimer (2012) 'Intelligence: New findings and theoretical developments', *American Psychologist*, 67(2): 130–59.
- Noakes, C. (2011) 'The historical and philosophical origins of normativism', *Behavioral and Brain Sciences*, 34(5): 253–4.

- Norman, D. and T. Shallice (1986) 'Attention to action: Willed and automatic control of behavior', in R. J. Davidson, G. E. Schwartz and D. Shapiro (eds), *Consciousness and Self-Regulation: Advances in Research and Theory* (vol. 4), New York: Plenum, 1–18.
- Norman, D. A. (1981) 'Categorization of action slips', *Psychological Review*, 88: 1–15.
- Nottelmann, N. (2006) 'The analogy argument for doxastic voluntarism', *Philosophical Studies*, 131: 559–82.
- Nottelmann, N. (2008) 'The present and future state of epistemic deontology', in V. Hendricks and D. Pritchard (eds), *New Waves in Epistemology*, Basingstoke: Palgrave Macmillan, 75–105.
- Nottelmann, N. (2013) 'The deontological conception of epistemic justification: A reassessment', *Synthese*, 190(12): 2219–41.
- Nozick, R. (1981) *Philosophical Explanations*, Cambridge, MA: Harvard University Press.
- O'Connor, T. (1994) 'Emergent properties', *American Philosophical Quarterly*, 31(2): 91–104.
- O'Connor, T. (1995) 'Agent causation', in T. O'Connor (ed), *Agents, Causes, and Events: Essays on Indeterminism and Free Will*, New York: Oxford University Press, 173–200.
- O'Connor, T. (2000) *Persons & Causes: The Metaphysics of Free Will*, Oxford: Oxford University Press.
- O'Connor, T. (2011) 'Agent causal theories of freedom', in R. Kane (ed), *The Oxford Handbook of Free Will*, 2nd edn, New York: Oxford University Press, 309–28.
- O'Keefe, T. (2010) *Epicureanism*, Durham: Acumen.
- Oaksford, M. and N. Chater (1992) 'Bounded rationality in taking risks and drawing inferences', *Theory & Psychology*, 2: 225–30.
- Otsuka, M. (1998) 'Incompatibilism and the avoidability of blame', *Ethics*, 108: 685–701.
- Owens, D. (2000) *Reason without Freedom*, London: Routledge.
- Oxford English Dictionary (1989) 2nd edn, Oxford, Clarendon Press.
- Paton, H. J. (1965) *The Categorical Imperative*, 5th edn, London: Hutchinson.
- Pereboom, D. (2001) *Living Without Free Will*, Cambridge: Cambridge University Press.
- Pereboom, D. (2011) 'Free will, skepticism and meaning in life', in R. Kane (ed), *Oxford Handbook of Free Will*, 2nd edn, New York: Oxford University Press, 407–24.
- Pettit, P. and G. Brennan (1986) 'Restrictive consequentialism', *Australasian Journal of Philosophy*, 64(4): 438–55.
- Piaget, J. (1952) 'Autobiography', in E. G. Boring (ed), *A History of Psychology in Autobiography*, New York: Russell & Russell, 237–56.
- Piguet, O., G. Grayson, A. Browe, H. Tate, T. Lye and H. Creasey (2002) 'Normal aging and executive functions in "old-old" community dwellers: Poor performance is not an inevitable outcome', *International Psychogeriatric Association*, 14: 139–59.
- Plantinga, A. (1993) *Warrant: the Current Debate*, Oxford: Oxford University Press.
- Plato (1956) *Meno*, tr. W. K. C. Guthrie, Harmondsworth: Penguin.
- Plato (1992) *Theaetetus*, tr. Levett (revised Burnyeat), Indiana: Hackett.
- Popper, K. (1966) *Of Clouds & Clocks*, St Louis: Washington University Press.

- Popper, K. (1982) *The Open Universe: An Argument for Indeterminism*, London: Routledge.
- Popper, K. (1983) 'Is determinism self-refuting?', *Mind*, 92: 103–4.
- Popper, K. and J. Eccles (1977) *The Self and Its Brain*, Berlin, Heidelberg, New York: Springer-Verlag.
- Prior, A. N. (1957) 'Modal and deontic logic', appendix D of *Time and Modality*, Oxford: Clarendon Press.
- Purves, D., R. Cabeza, S. Huettel, K. LaBar, M. Platt and M. Woldorff (2013) *Principles of Cognitive Neuroscience*, 2nd edn, Sunderland, MA: Sinauer.
- Pust, J. (2001) 'Against explanationist skepticism regarding philosophical intuitions', *Philosophical Studies*, 106: 227–58.
- Putnam, H. (1975) 'The nature of mental states', in *Mind, Language and Reality Philosophical Papers* (vol. 2), Cambridge: Cambridge University Press, 429–40.
- Putnam, H. (1980) 'Philosophy and our mental life', in N. Block (ed), *Readings in Philosophy of Psychology*, Cambridge: Harvard University Press, 134–43.
- Putnam, H. (1983) 'Philosophers and human understanding', in *Realism & Reason: Philosophical Papers* (vol. 3), Cambridge: Cambridge University Press, 184–204.
- Quine, W. V. (1948/1980) 'On What There Is' in his *From a Logical Point of View*, 2nd edn, Cambridge, MA: Harvard University Press, 1–19.
- Reason, J. T. (1992) 'Cognitive underspecification: its variety and consequences', in B. J. Baars (ed), *Experimental Slips and Human Error: Exploring the Architecture of Volition*, New York: Plenum Press, 71–91.
- Reid, T. (1969) *Essays on the Active Powers of the Human Mind*, Cambridge, MA: MIT Press.
- Rips, L. (1983) 'Cognitive processes in propositional reasoning', *Psychological Review*, 90: 38–71.
- Rudder Baker, L. (1988) *Saving Belief: a Critique of Physicalism*, Princeton: Princeton University Press.
- Rudder Baker, L. (2006) 'Moral responsibility without libertarianism', *Nous*, 40(4): 307–30.
- Russell, B. (1996) *A History of Western Philosophy*, London: Routledge.
- Russell, P. (2008) 'Free will, art and morality', *Journal of Ethics*, 12: 307–25.
- Russell, P. (2010) 'Selective hard compatibilism', in J. Campbell, M. O'Rourke and H. Silverstein (eds), *Action, Ethics and Responsibility: Topics in Contemporary Philosophy* (vol. 7), Cambridge, MA: MIT Press, 149–74.
- Ryan, S. (2003) 'Doxastic compatibilism and the ethics of belief', *Philosophical Studies*, 114: 47–79.
- Ryding, E., B. Bradvik and D. Ingvar (1996) 'Silent speech activates prefrontal cortical regions asymmetrically, as well as speech-related areas in the dominant hemisphere', *Brain & Language*, 52: 435–51.
- Sartwell, C. (1991) 'Knowledge is merely true belief', *American Philosophical Quarterly*, 28(2): 157–65.



- Sartwell, C. (1992) 'Why knowledge is merely true belief', *Journal of Philosophy*, 89(4): 167–80.
- Schmahmann, J. (2004) 'Disorders of the cerebellum: ataxia, dysmetria of thought, and the cerebellar cognitive affective syndrome', *Journal of Neuropsychiatry & Clinical Neuroscience*, 16(3): 367–78.
- Schopenhauer, A. (1966) *The World as Will and Representation Vol II*, tr. E. F. J. Payne, New York: Dover Publications.
- Schopenhauer, A. (1999) *Prize Essay on Freedom of the Will*, tr. E. F. J. Payne, Cambridge: Cambridge University Press.
- Sechenov I. M. (1942) *Refleksy golovnogo mozga*; cited in M. U. Barbashin (2008) 'Methodological Foundations of Ethnopolitical Conflict Studies', *The Peace & Conflict Review* 2(1) issn 1659-3995, available: <http://www.review.upeace.org/pdf.cfm?articulo=74&ejemplar=14> (accessed 9 June 2017).
- Seligman, M., P. Railton, R. Baumeister and C. Sripada (2013) 'Navigating into the future or driven by the past', *Perspectives on Psychological Science*, 8(2): 119–41.
- Sergeant, J., H. Geurts and J. Oosterlaan (2002) 'How specific is a deficit of executive functioning for attention-deficit/hyperactivity disorder?', *Behavioural Brain Research*, 130: 3–28.
- Shabo, S. (2010a) 'The fate of the direct argument and the case for incompatibilism', *Philosophical Studies*, 150(3): 405–24.
- Shabo, S. (2010b) 'Uncompromising source incompatibilism', *Philosophy and Phenomenological Research*, 80(2): 349–83.
- Shabo, S. (2011) 'What must a proof of incompatibilism prove?', *Philosophical Studies*, 154(3), 361–71.
- Shallice, T. (1982) 'Specific impairments of planning', *Philosophical Transactions of the Royal Society of London; Series B, Biological Sciences*, 298(1089): 199–209.
- Shatz, D. (1988) 'Compatibilism, values, and "could have done otherwise"', *Philosophical Topics*, 16(1): 151–200.
- Shepard, R. N. and J. Metzler (1971) 'Mental rotation of three-dimensional objects', *Science* 171(3972): 701–3.
- Sidgwick, H. (1907) Appendix to *The Methods of Ethics*, London: Macmillan; revised from first publication in *Mind*, 13(51): 1888.
- Simon, H. and W. Chase (1973) 'Skill in chess', *American Scientist*, 61: 393–403.
- Skinner, B. F. (1971) *Beyond Freedom and Dignity*, New York: Knopf.
- Skrubina, D. (2005) *Panpsychism in the West*, Cambridge, MA: Bradford Books: MIT Press.
- Smart, J. J. C. (1959) 'Sensations and brain processes', *Philosophical Review*, 68: 141–56.
- Smart, J. J. C. (1961) 'Free will, praise and blame', *Mind*, 70: 291–306.
- Smith, E. and S. Kosslyn (2007) *Cognitive Psychology*, Upper Saddle River: Pearson.
- Smith, H. M. (1988) 'Making moral decisions', *Nous*, 22: 89–108.
- Sperry, R. W. (1986) 'Macro- versus micro-determinism', *Philosophy of Science*, 53: 265–70.

- Stanovich, K. and R. West (2000) 'Individual differences in reasoning: Implications for the rationality debate?' *Behavioral and Brain Sciences*, 23: 645–65.
- Sternberg, R. J. (2009) *Cognitive Psychology*, 5th, International Student, edn, Wadsworth: Cengage.
- Steup, M. (2008) 'Doxastic freedom', *Synthese*, 161(3): 375–92.
- Steward, H. (2012) *A Metaphysics for Freedom*, Oxford: Oxford University Press.
- Stich, S. (1990) *The Fragmentation of Reason*, Cambridge, MA: Bradford Books: MIT Press.
- Stich, S. (1994) 'Could man be an irrational animal?', in H. Kornblith (ed), *Naturalizing Epistemology*, 2nd edn, Cambridge, MA: Bradford Books: MIT Press.
- Strathman, A., F. Gleicher, D. Boninger and C. Edwards (1994) 'The consideration of future consequences: Weighing immediate and distant outcomes of behavior', *Journal of Personality and Social Psychology*, 66(4): 742–52.
- Strawson, G. (1995) 'Libertarianism, action, and self-determination', in T. O'Connor (ed), *Agents, Causes, and Events: Essays on Indeterminism and Free Will*, Oxford: Oxford University Press, 13–31.
- Strawson, G. (2003) 'Mental ballistics: The involuntariness of spontaneity', *Proceedings of the Aristotelian Society*, 103(1): 227–56.
- Strawson, P. F. (1962) 'Freedom and resentment', *Proceedings of the British Academy*, 48: 1–25.
- Stump, E. (2003) 'Moral responsibility and alternative possibilities', in D. Widerker and M. McKenna (eds), *Freedom, Responsibility and Agency: Essays on the Importance of Alternative Possibilities*, Aldershot: Ashgate Press: 139–58.
- Thomas, L. and A. Lleras A. (2007) 'Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition', *Psychonomic Bulletin & Review*, 14(4): 663–8.
- Thomas, L. and A. Lleras (2009) 'Swinging into thought: Directed movement guides insight in problem solving', *Psychonomic Bulletin & Review*, 16(4): 719–23.
- Timpe, K. (2007) 'Source incompatibilism and its alternatives', *American Philosophical Quarterly*, 44(2): 143–55.
- Tolman, E. C. (1948) 'Cognitive maps in rats and men', *Psychological Review*, 55(4): 189–208.
- Tolstoy, L. (1983) 'Postface to the Kreutzer Sonata', in *The Kreutzer Sonata and other Stories*, tr. D. McDuff, Harmondsworth: Penguin, 267–82.
- Van Inwagen, P. (1980) 'The incompatibility of responsibility and determinism', in M. Bradie and M. Brand (ed), *Bowling Green Studies in Applied Philosophy Vol 2: Action and responsibility*, Bowling Green: Bowling Green State University Press, 30–7.
- Van Inwagen, P. (1983) *An Essay on Free Will*, Oxford: Oxford University Press.
- Van Inwagen, P. (2000) 'Free will remains a mystery', *Philosophical Perspectives*, 14: 1–20.
- Vygotsky, L. (1962) *Thought & Language*, ed. and tr. E. Hanjmann and G. Vakar, Cambridge, MA: MIT Press.

- Vygotsky, L. (1978) *Mind in Society*, tr. M. Cole, Cambridge, MA: Harvard University Press.
- Warfield, T. A. (2000) 'A new argument for incompatibilism', *Philosophical Perspectives*, 14: 167–80.
- Watson, G. (1975) 'Free agency', *Journal of Philosophy*, 72(8): 205–20; reprinted in G. Watson (ed), *Free Will*, Oxford Readings in Philosophy, 2nd edn, Oxford: Oxford University Press, 2003, 337–51.
- Waugh, E. (1960) *Brideshead Revisited: The Sacred and Profane Memories of Captain Charles Ryder*, London: Chapman & Hall.
- Weatherson, B. (2008) 'Deontology and descartes' demon', *Journal of Philosophy*, 105: 540–69.
- Weinberg, J. M. (2006) 'What's epistemology for? The case for neopragmatism in normative metaepistemology', in S. Hetherington (ed), *Epistemology Futures*, Oxford: Oxford University Press, 26–47.
- Wendell Holmes, O. (1992) *Selections from the Letters, Speeches, Judicial Opinions, and Other Writings of Oliver Wendell Holmes, Jr.*, R. A. Posner (ed), Chicago: Chicago University Press.
- Widerker, D. (1991) 'Frankfurt on "ought" implies "can" and alternative possibilities', *Analysis*, 49: 222–4.
- Widerker, D. (1995) 'Libertarianism and Frankfurt's attack on the principle of alternative possibilities', *Philosophical Review*, 104: 247–61.
- Widerker, D. (2000) 'Frankfurt's attack on the principle of alternative possibilities: A further look', *Philosophical Perspectives*, 14: 181–201.
- Widerker, D. (2002) 'Farewell to the direct argument', *Journal of Philosophy*, 99: 316–24.
- Widerker, D. (2009) 'A defense of Frankfurt-friendly libertarianism', *Philosophical Explorations*, 12(2): 87–108.
- Widerker, D. (2011) 'Frankfurt-friendly libertarianism', in R. Kane (ed), *Oxford Handbook of Free Will*, 2nd edn, New York: Oxford University Press, 266–87.
- Willcutt, E., A. Doyle, J. Nigg, S. Faraone and B. F. Pennington (2005) 'Validity of the executive function theory of attention deficit/hyperactivity disorder: A meta-analytic review', *Biological Psychiatry*, 57: 1336–46.
- Williamson, T. (2007) *The Philosophy of Philosophy*, Oxford: Oxford University Press.
- Wilson, R. (2004) *Boundaries of the Mind: The Individual in the Fragile Sciences – Cognition*, Cambridge: Cambridge University Press.
- Wilson, R. and A. Clark (2009) 'How to situate cognition: Letting nature take its course', in P. Robbins and M. Aydede (eds), *The Cambridge Handbook of Situated Cognition*, Cambridge: Cambridge University Press, 55–77.
- Windelband, W. (1894/1980) 'Rectorial address, Strasbourg, 1894', tr. G. Oakes, *History and Theory*, 19(2): 169–85.
- Wittgenstein, L. (1953) *Philosophical Investigations*, tr. Anscombe, Oxford: Blackwell.
- Wittgenstein, L. (1958) *The Blue and Brown Books*, Oxford: Basil Blackwell.
- Wolf, S. (1990) *Freedom within Reason*, Oxford: Oxford University Press.

- Wright, C. (2002) '(Anti-)sceptics, simple and subtle: G.E Moore and John McDowell', *Philosophy and Phenomenological Research*, 65: 330–48.
- Wu, W. (2014) *Attention*, Oxford: Routledge.
- Yablo, S. (1992) 'Mental causation', *Philosophical Review*, 101: 245–80.
- Yaffe, G. (2005) 'More on "ought" implies "can" and the principle of alternate possibilities', *Midwest Studies in Philosophy*, 29: 307–12.
- Zagzebski, L. (1996) *Virtues of the Mind*, Cambridge: Cambridge University Press.
- Zagzebski, L. (2003) 'The search for the source of epistemic good', *Metaphilosophy*, 34(1–2): 12–28.