Explicating agency: the case of visual attention[1]

Abstract: How do individuals guide their activities towards some goal? Harry Frankfurt once identified the task of explaining guidance as the central problem in action theory. An explanation has proved to be elusive, however. In this paper I show how we can marshal empirical research to make explanatory progress. I contend that human agents have a primitive capacity to guide visual attention, and that this capacity is actually constituted by a sub-individual psychological control-system: the executive system. I thus illustrate how we can explain exercises of individual-level guidance by appeal to its sub-individual constitution. This opens up a new avenue for explicating agency.

Words: 9466

Keywords: action, guidance, attention, executive control

1 Explicating agency

Actions most fundamentally are an individual's doings. When we act, we make something happen, rather than merely undergoing it. In doing something, an agent is active, throughout the action's execution. What is it for an individual to act?

The most influential answer to this question is provided by causal theories of action. This family of views has it that an action is an event, appropriately caused by the right kind of mental antecedent. Much work has gone into determining what the right kind of mental antecedent might be: different theorists have proposed desires, belief/desire-pairs, or intentions as candidates, to name just a few. (Davidson 1970; Searle 1983; Brand 1984; Bratman 1987; Bishop 1989) The view's focus is on how an agent's mind causes some occurrence in the world. From the outset, the view treated an action's execution as an afterthought.

But surely, our answer is at best incomplete, if it does not explain how an agent is doing something, throughout an action's execution. This observation was first made by Harry Frankfurt. He contends that the causal theory's focus on action's mental antecedent makes it

> "impossible for [it] to give any account whatever of the most salient differentiating characteristic of action: during the time a person is performing an

[1] Special thanks to Tyler Burge. Thanks also to Ned Block, Susan Carey, Martin Davies, Harry G. Frankfurt, Pam Hieronymi, Kevin Lande, Al Mele, Elisabeth Pacherie, Chris Peacocke, Ian Phillips, Thor Gruenbaum, Hong Yu Wong, and Wayne Wu. I acknowledge funding from ANR-17-EURE-0017.

> action he is necessarily in touch with the movements of his body in a certain way, whereas he is necessarily not in touch with them in that way, when the movements of his body are occurring without his making them." (Frankfurt 1978, 71)

Frankfurt identifies, as the target explanandum for a theory of action, not its mental antecedent causes, but rather

> "whether or not the movements as they occur are under the person's guidance. It is this that determines whether he is performing an action." (Frankfurt 1978, 72)

Frankfurt assumes, with the causal theory, that all action is purposive. But, he observes, so are many other processes, such as the growth of a plant, or the dilation of our pupils, or the transport of nutrients in our blood stream. We might think of these processes are fulfilling biological functions, or as being directed toward some end-state. But in those cases, no human agent does the directing. For Frankfurt, action is behavior "whose course is under the guidance of an agent." (Frankfurt 1978, 74) And accordingly, for him, explicating agency requires explaining guidance. He writes:

> "An explication of the nature of action must deal with two distinct problems. One is to explain the notion of guided behavior. The other is to specify when the guidance of behavior is attributable to an agent, and not simply, as when a person's pupils dilate because the light fades, to some local process going on within the agent's body." (Frankfurt 1978, 74)

I believe that Frankfurt is fundamentally right. To explicate agency, we must explain in virtue of what an agent is doing something, while she is doing it. In goal-directed action, we must explain in virtue of what it is the agent that is directing her activities toward some goal. We must explain what it is for an individual to actively guide her activities, throughout the action's execution. Providing such an explanation constitutes the *problem of guidance*.

Causal theorists, of course, have been sensitive to this issue. But, arguably, no causalist treatment of it has been adequate. One strand within the causalist tradition identifies Frankfurt's problem as that of the "disappearing agent."[2] Causal theories

[2] Very different problems have been called the "disappearing agent" problem. Some versions have worried about conditions on free agency. (Pereboom 2004) Others have worried about the compatibility of agency with a metaphysics of causation in terms of events and states. (Velleman 1992; Hornsby 1996) Yet other

attempt to explain individuals' agency in terms of (psychological) states, events, and causal relations between them. But, goes this problem, agents are not identical to (some of) their states and events; they rather are their bearers. The worry, as Velleman puts it, is that on the causal theory,

> "nobody – that is, no person – does anything. Psychological and physiological events take place inside a person, but the person serves merely as the arena for these events: he takes no active part." (Velleman 1992, 461)

The solution, for these theorists, is that of finding

> "mental events and states that are functionally identical to the agent, in the sense that they play the causal role that ordinary parlance attributes to him." (Velleman 1992, 475)

Desires, belief/desire-pairs, and intentions, for these theorists, just are not the right kind of mental state or event to explain how the agent herself guides her action. For Velleman, it is the higher-order desire to act in accordance with reason that plays this role. (Velleman 1992, 479) Other theorists have proposed other solutions. (Frankfurt 1988; Bratman 2007; Korsgaard 2009)

But this kind of solution is hopeless as an answer to the problem of guidance. Not only does it over-intellectualize human agency: all human action, on this picture, must be driven by some higher-order, reflective state, representing the kind of first-order state that the agent wants her actions to be caused by. The reflective state is said to involve recognition of one's reasons, intentions, or policies as such. Surely, much human agency is not so sophisticated.

Even apart from this observation, the view does not even begin to explain how it is that the individual guides an action, throughout its execution. The view merely stipulates yet another mental antecedent cause for action. While this type of view seems right to emphasize questions about the agent's proper role in action, it entirely misses Frankfurt's emphasis on explaining that role *during* the action's *execution*.

---

versions have worried about the place for agents in a scientific world view. (Nagel 1986; Bishop 1989; Hornsby 2004; Steward 2012)

Another strand within the causalist tradition takes Frankfurt to be raising problems of causal deviance. Canonical statements of a causal theory have it that the relevant mental antecedent must cause *in the right way* the event that constitutes the action. The challenge is to specify sufficient conditions for such causation to constitute action. As Bishop puts it:

> "[F]or CTA to be proved, we need to establish a CTA analysis that specifies purely event-causal conditions that are *sufficient* as well as necessary for the occurrence of action." (Bishop 1989, 101)

An intention to kill your fellow mountaineer by letting go of the rope that holds him, might cause you to become nervous, to unintentionally loosen your grip, and so to kill your partner. In this case, a relevant mental antecendent plays a causal role in bringing about the intended event. But this kind of causal chain does not constitute action. If we knew how to rule out deviance, then, arguably, this would allow us to explain how it is that an agent is guiding her activities, during an action's execution.[3] Unfortunately, the challenge from deviance is considered to be wide open.[4]

Some of the most sustained and sophisticated replies have sketched causal mechanisms that must be operative in non-deviant action. These mechanisms' operation is said to be causally sustained by an intention that persists throughout the action's execution. (Mele 1992) The mechanisms monitor the behavior's progress and its congruence with the intention's content. (Bishop 1989; Adams & Mele 1989; Mele 1992) But not only has each reply faced counter-examples and criticisms.

It is unclear what shape a satisfying answer might take. The debate about deviance has taught us that not any causal mechanism that brings about intended behavior constitutes action. Even in actual human agents, a plethora of causal mechanisms are operative at any given moment, and during each exercise of agency. Which of these is the right mechanism to constitute the agent's guidance? We must identify a causal mechanism that helps constitute the agent – as she regulates her activities, *throughout* the execution of her action. (Velleman 1992) But what does such a mechanism have to look

[3] I do not think, however, that we need to specify such sufficient conditions, in order to explicate individuals' guidance.
[4] A range of recent attempts have been made to address the challenge. (Brent 2017; Wu 2016; Shepherd 2014; O'Brien 2012) Discussing them here would lead to far afield.

like? The main aim of this paper is to make progress on this last question: how to argue that some proposed mental state or mechanism constitutes the *individual's* active guidance. We cannot satisfyingly resolve challenges from disappearing agents or deviance, without first addressing this question, I believe.

Action theory has largely proceeded from the armchair, by addressing ever more elaborate counter-examples. In this paper I want to explore a different approach to the problem of guidance. This approach starts from ostension. It takes as its starting point our best empirical theories of goal-directed action. Having identified what constitutes an actual capacity to guide in human agents, we can use this to direct our reflection on the kind *human agency*. This general type of strategy has led to important insights in the philosophy of perception. (Burge 2010) And in any case, the fact that there has been little progress on the problem of guidance suggests that a novel approach might be worth a try.

I will argue that actual human agents have at least one capacity to actively guide: a capacity to guide visual attention shifts. This capacity, I will further argue, is actually constituted by an empirically discovered system for cognitive control – the executive system. Appeals to this capacity-constituting sub-system, I further claim, allow us to explain individuals' guidance, and through it, to begin to explicate human agency. I focus on active shifts of visual attention, not only because this is an especially important kind of action, often overlooked by action theorists: a basic form of mental action, central to inquiry and acting full well. This kind of action is also particularly well understood in cognitive science, and hence lends itself to the kind of investigation I propose.

Note two important qualifications. First, traditional action theory is interested in the nature of action. Reflection on the actual human capacity to act does not, in any straightforward way, yield insight into action's nature. That is why I make no modal claim. I focus on actual human agents. Expanding the claim will require further argument. Second, my proposal focuses on one kind of guided activity: active shifts of visual attention. For this type of human action, I argue that the executive system constitutes a capacity to guide. I do here not address other kinds of human action. What I propose is a partial empirical explication of agency. (Carnap 1950) While I do believe that the claim holds across all the varieties of human agency, I do not argue this claim here. I am rather concerned with laying the groundwork for a more general argument.

A subsidiary aim of the paper is thus to explore how we can use empirical insight to explicate agency. How might one go about this? The first step must be a philosophical characterization of the explanandum, as well as an analysis of extant empirical work, to motivate a conjecture for a candidate-explanans. This happens in sections 2-4, where I circumscribe properties of individuals' capacity to guide, introduce the executive system, and show that this executive system operates in attention-shifts that are guided by the individual, but not in others. The conjecture raises a range of philosophical issues: How can we establish that this executive system not merely correlates with, but *constitutes* a capacity to guide visual attention? What kind of explication of agency does this provide us with, if it does not give us the essence of agency? And how can such an explication be neither vacuously circular, nor explain away a role for the individual in action's execution? In section 5 I describe a conception of capacity-constitution that allows us to address each of these questions. Sections 6 and & 7 discuss worries and alternative proposals. In section 8 I conclude by reflecting on where this leaves the project of explicating agency.

2 A capacity to guide

What characterizes individuals' putative capacity to guide visual attention? Suppose that you are a birder, out in the woods. You try to spot a robin. You have a rough mental image of its looks: its shape, coloring, and peculiarities of movement. You search the crowns by shifting attention – and with it, often, your eyes – across the foliage. You first sample the scene with a few large saccades. Next you more narrowly scan areas that most likely contain birds. Your eyes move to stimuli that resemble the bird. Often, this resemblance misleads. You spot another bird, or animal, or an oddly shaped leaf. Sometimes, you look at things merely because they stand out, because they are oddly colored or weirdly shaped. But you adjust: you avoid already visited locations and get better at ignoring distractions. You search the scene until you finally find the robin.

Such a birder *guides* her attention-shifting activity toward her goal, in search for the robin. We can contrast such guided attention-shifts with cases in which your attention is captured. Suppose that, while you are looking for the robin, you suddenly hear a loud bang from right behind you. Or suppose that a snake-shaped object appears in the corner of your eye. Or maybe you see a bright flash. These events will disrupt your search for the robin. They force you to attend to them, instead. In these cases, attention shifts passively. You do not guide attention toward those stimuli.

Central instances of guidance are characterized by a set of surface-properties, or marks. These marks do not constitute a definition, but they circumscribe, in a preliminary, defeasible way, our target-explanandum. Frankfurt writes that guided behavior is purposive, or goal-directed. The course of such behavior

> "is subject to adjustments which compensate for the effects of forces which would otherwise interfere with the course of the behavior." (Frankfurt 1978, 74)

Others have emphasized that such action involves a kind of

> "functioning, coordinated behavior by the whole organism … [that] must issue from central capacities, in effect coordinating sub-systems, or coordinating central capacities with their peripheral realizations." (Burge 2009, 260)[5]

Taking a cue from these theorists, we can characterize our explanandum as follows. When an individual guides, roughly, the activities of her parts are flexibly coordinated, so as to be *directed toward her goal*. Such *coordination* has both synchronic and diachronic aspects. The birder may hold active an image of the bird, while she searches for its likeness in the trees, and simultaneously keep her posture still. The birder may first search the crowns, then the branches closer to the ground. In so coordinating, the birder relies on information *integrated* from a range of sources. The birder integrates proprioception of her body's position with auditory information about the direction of the bird's song. She integrates visual information with memories about the bird's likely location to direct her attention. Finally, the birder coordinates flexibly: she *compensates* for interference and adjusts activities in light of incoming information. She may ignore a distracting bright light when shifting attention from one location to the next. Or she may suppress an urge to stretch her legs, to avoid frightening the animal. Importantly, when individuals guide, they *make* this flexible coordination of their activities. But their

[5] We find similar ideas in Steward 2012 and Hyman 2015.

coordinating, integrating, and compensating, in basic cases, *is not itself a further act*. It rather helps constitute the individual's action.

I suggest that these surface-properties, or marks, are due to the exercise of a putative capacity to guide shifts of visual attention. This capacity typically generates processes that exhibit synchronic and diachronic coordination, that rely on information-integration from a wide range of sources, and that flexibly compensate for interference. None of these properties are necessary for an exercise of guidance. Nor are they individually or jointly sufficient. They rather help us provide a first-pass characterization of the capacity at issue. The suggestion raises legitimate questions: Is there *really* such a primitive capacity? How, and on what grounds, can we further characterize this capacity? How can we explain its operation in an illuminating, non-circular way? It is here that I propose we turn to a philosophical analysis of pertinent work in empirical psychology.

3 The executive system

In the next three sections I will argue that our capacity to guide visual attention is actually constituted by the executive system.[6] Appeal to this sub-system allows us to answer the questions raised earlier. As a first approximation, the executive system is an empirically discovered psychological sub-system for the control of processes in other psychological sub-systems. Individuals' psychologies are roughly hierarchically structured, containing a range of sub-systems devoted to ever more specialized tasks, including perceptual, memory, and motor systems. The executive system accesses and regulates a wide range of these sub-systems. The system organizes activity in these sub-systems for carrying out the whole individual's tasks.

The system regulates other sub-systems through the exercise of four control-competencies: the executive functions of switching, maintenance, resource-allocation, and inhibition. *Executive switching* activates the suite of representations and abilities

[6] My conception of this system *derives* from the specific strand of research documented in the main text. I do not commit to details of specific scientific proposals. Different notions of 'executive systems' have been used in the literature. Not all of them are equivalent to what I sketch in the main text.

pertinent to carrying out some task (sometimes called the “task set”).[7] The mechanism underlying switching determines which representations and abilities are relevant to some task, initiates, and fixes parameters for such processing in light of the individual’s goal. *Executive maintenance* encodes information into working memory and maintains it active during the execution of a task. Such maintenance relies on mechanisms for determining which information is relevant for the task at hand. *Executive resource-allocation* involves the deployment of executive processing resources for the execution of a task.[8] These resources function to enhance the processing to which they are allocated. A mechanism determines what processes to flexibly allocate them to. *Executive inhibition*, finally, is a competence for suppressing the influence of distractors or prepotent responses on ongoing activity. It relies on mechanisms that determine processing of which stimuli or impulses constitutes interference with the current task, and that suppress processing of such stimuli.

Appeal to these four competencies serves to characterize the executive system, to identify its operations, its contribution to processes, and to guide theorizing about it. They are signatures of this system. (Miyake et al. 2000; Miller & Cohen 2001; Baddeley 2007; Diamond 2013; Goldstein et al. 2014; Botvinick & Cohen 2014; Gazzaniga 2014; Fuster 2015)

The competencies typically cooperate to control processes. They access and regulate processing in other relevant sub-systems, in light of information about the individual’s goal. The system, one might say, oversees and steers independent processing in other sub-systems. But not all four competencies must be engaged for the executive system to exert control.

Characterizations of these competencies and their interaction offer a cognitive model of the executive system. (Weiskopf 2018) This model derives from behavioral psychology. Executive competencies were initially studied through behavioral paradigms that were assumed to fairly exclusively require the exercise of only one executive function, such as the Antisaccade, Stroop, or Simon test. (Miyake et al. 2000; Diamond

[7] Typically, such activation requires de-activation of the set for whatever activity the individual carried out prior, hence “switching” from one set to another. (Sakai 2008)
[8] Sometimes this executive function is called “executive attention.” This label may be misleading. Attention is an individual-level capacity. Executive resource-allocation need not be.

2013) Subjects would exhibit characteristic patterns of performance at these tests: specific time-courses, error-rates or error-patterns, and effects of intervention on aspects of the tests. Characteristic patterns of performance were explained in terms of the underlying exercise of executive functions. The executive functions in turn were characterized by their modes of operation, the kinds of representations that they operated on, and the resources that they had available for their operation. In this way, researchers obtained initial characterizations of the executive functions. These initial characterizations then generated more specific hypotheses about executive functioning in novel experimental contexts, as well as about its interaction with other competencies. More specific hypotheses in turn led to further sharpening of the executive functions' characterization.[9]

The cognitive model of the executive system thus became ever more refined over the course of the years. Its sophistication increased as the executive functions' characterizations, and paradigms for studying them, sharpened. These advances in turn allowed increasing integration of behavioral research with neuroscientific studies and computational modeling. (Badre 2020)

Thus we now know that the brain roughly mirrors the hierarchy of psychological sub-systems. While sub-systems devoted to, for instance, motor or perceptual processing are distributed across posterior brain areas, frontal and especially prefrontal cortex appear to implement control structures. They modulate processing in other parts of the brain through long-range connections with those areas. This interconnectivity with almost all parts of the brain situates (pre-)frontal cortex especially well for regulating processing across the brain "in light of" instructions issued by (pre-)frontal cortex. (Miller & Buschman 2013) Indeed, it has been possible to identify specific parts of frontal cortex as implementing specific executive functions. (Duque et al. 2013; Gazzaniga 2014; D'Esposito & Postle 2015)

[9] Milestones of this behavioral research include work on cognitive control by Posner & Snyder 1975, Shiffrin & Schneider 1977, Norman & Shallice 1986, and Baddeley's model of working memory (Baddeley 1986, 2007). In this early work, researchers sometimes vaguely appealed to executive attention as explaining subjects' performance. Such appeals were, at the time, rightly criticized as not explanatory. But especially the last two decades have seen major advances in the rigor of this research. (Cf. Miyake et al. 2000; Diamond 2013)

Similarly, computational models have confirmed general assumptions about the executive functions and their characteristics. Thus models lacking a control network including basic executive functions will not approximate actual human performance. This result supports the idea that successful execution of even fairly simple behaviors requires executive functioning. (Rougier et al. 2005; Botvinick & Cohen 2014) Furthermore, it has been increasingly possible to design computational models of specific (aspects of) executive functions. Thus Koechlin et al. (2007) offered computational models of how specific neural networks implement computations surrounding executive switching. Ventromedial and dorsomedial regions, for instance, monitor the reliability of current set. Polar and lateral regions determine the relative reliability of several alternative task sets. These two networks interact to determine when to switch to what set. (Hyafil et al. 2009; Koechlin 2014; Donoso et al. 2014)

Findings from across different disciplines and spanning different explanatory levels and approaches – psychological, neuroscientific, computational modeling – present us with a detailed, sophisticated cognitive model, even a partial mechanism sketch, of the executive system and the ways in which it accesses and regulates processing across the hierarchy of psychological sub-systems. In what follows I will rely on this model.[10]

4 Orienting attention

We have encountered our target-explanandum – the putative capacity to guide visual attention. I have introduced the executive system, a sub-system for cognitive control that, I will argue, constitutes this capacity. In this section, I want to motivate that this system is a plausible candidate for the job. To do so, I will describe the role of the system in orienting visual attention. This requires analyzing a fair amount of empirical detail about visual attention. I will rely on some of this detail for my arguments in section

[10] It has sometimes been said that appeals to executive control are explanatorily vacuous. This worry may have been legitimate thirty years ago. Actual explanatory practice in cognitive science refutes the worry. Appeals to executive functioning in current cognitive science are not explanatorily empty. The executive functions are sufficiently independently characterized to offer explanations and novel, testable hypotheses that would not otherwise have been available. We will be able to more fully appreciate this fact once we have seen, in the following sections, how such explanation works. (Dennett 1978; Allport 1993; Monsell & Driver 2000; Buehler 2017)

5. Readers less interested in this detail may want to skim this section and return to it later, as needed.

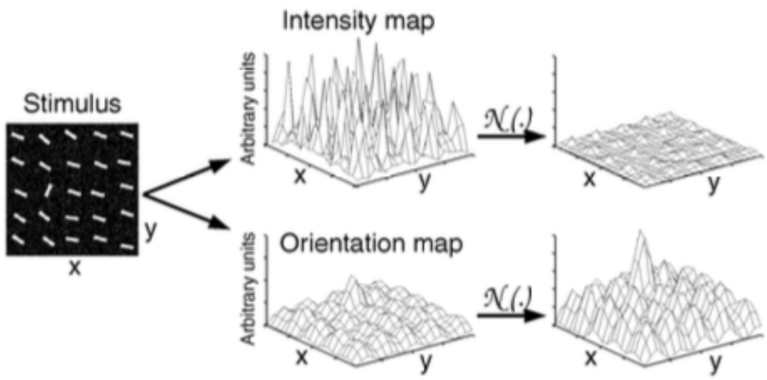
How does psychology explain visual attention-shifts?[11] Psychology standardly distinguishes between two sub-systems involved in orienting visual attention: the exogenous and the endogenous system. The exogenous system primarily reacts to physically or practically salient stimuli, such as an abruptly appearing bright light. Roughly, saliency is a measure of the extent to which a stimulus stands out from its immediate surroundings. Stimuli with high practical relevance might be predators or mates, for example. The exogenous system shifts attention to the salient stimulus rapidly, within only 120 ms. The system affects visual processing of that stimulus in characteristic ways. Where attention shifts exogenously is largely independent of individuals' current goals and expectations. Indeed, individuals cannot suppress orientation to a salient stimulus even when they know that it is a distractor and *try to ignore it*. (Itti & Baldi 2009; Bruce & Tsotsos 2009; Theeuwes 2010)

The endogenous system takes individuals' goals and intentions as inputs. This system shifts attention in light of a goal to find a bird, or a decision to fixate some location on a wall. The system orients attention slowly, taking about 300 ms to do so. It, too, has characteristic effects on visual processing, but they differ from those found for exogenously oriented attention. In cases of deliberate fixation of attention, it appears that the individual's intention alone determines where attention shifts. Thus, (if given sufficient time) individuals can decide to fixate some location – and not be distracted by *even the most salient distractor*. (Theeuwes 1991; Yantis & Jonides 1990)

Both systems contribute to orienting attention by influencing priority-assignments for locations on a priority map. This map is a topographical representational structure that assigns priority-values to locations in the scene. (See Fig. 1) Neural structures in frontal eye fields, lateral intraparietal area, and superior colliculus implement the map.

[11] Several constitutive accounts of attention have been proposed over the last years. Attention has been said to consist in our mental capacities' cognitive unison (Mole 2011), the selection of a stimulus for a response by the individual (Wu 2014), the regulation of priority structures (Watzl 2017), the making-available of information to thought (Smithies 2011) and to working memory (Prinz 2012). I commit to the claims that visual attention involves the allocation of visual processing resources and can be shifted in the ways described in the main text. I otherwise remain non-committal as regards the nature of attention. What I say about visual attention is, as far as I can see, compatible with each of these accounts.

(Buschman & Kastner 2015; Zelinsky 2015) Attention shifts to the location with the highest priority-value on this map.

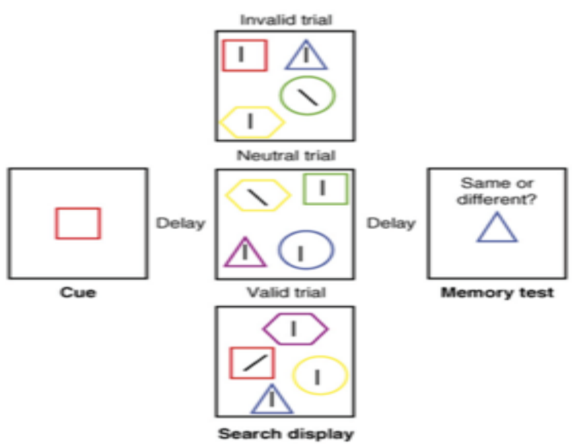


[Fig. 1: A bar yielding high priority on a map registering stimuli that stand out from their surroundings as regards their orientation. (Itti et al. 1998) The priority map partly results from pooling such saliency maps.]

Could we not identify the endogenous system as the one that constitutes individuals' capacity to guide visual attention, and take the exogenous system to be responsible for non-guided, passive, captured attention-shifts? No. Even in paradigmatically active, guided attention-shifts, the exogenous system sometimes positively contributes to the individual's guidance. On the other hand, components of the endogenous system can interfere with individuals' goal-directed guidance. Generic appeal to the endogenous system would not yield a plausible candidate for constituting our attention-guiding capacity.

Thus take a case of visual search, like that of the birder from section 2. In visual search, individuals set a goal-representation of a search-target, they initiate and guide attention-shifts in search of their goal. Often during visual search, salient or practically highly relevant stimuli, processed by the exogenous system, attract attention. Suppose that we instruct a subject to search for a green diamond-symbol in a display. Further suppose that all other items in the display are red diamonds. The target item, of course, is physically salient in the display. It stands out, due to its unique color. The exogenous system will assign high priority to its location. This assignment makes more likely that the individual will shift attention there and speed up the subject's search. Her guidance benefits from the exogenous system's operation. (Leonard & Egeth 2008) Despite its influence on her search, the individual still guides attention: it is her search-goal that drives the search, and she directs where attention goes.

On the other hand, components of the endogenous system, such as implicit or working-memories, can interfere with the individual's guidance of her search. Consider implicit memory due to priming, for example. If a subject repeatedly searches displays for green diamond shapes, then priority for stimuli of that type will be boosted by default. The subject will be faster and more accurate at detecting green diamond shapes. This effect lasts, even when her search goal changes. So even if now she attempts to find a red circle shape in a display, green diamond shapes will tend to attract her attention.[12] The effect will slow down her search, she will make more errors, and may shift attention to the green distractor. (Kristjanson & Campana 2010)[13]

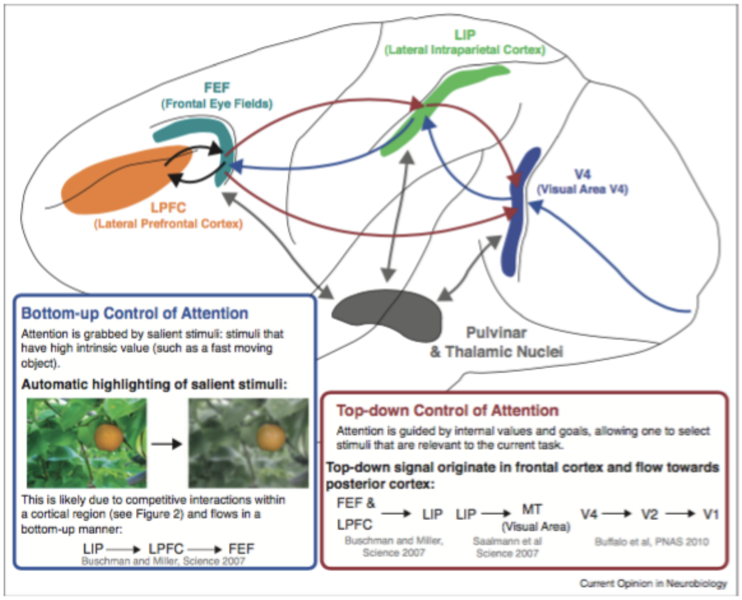


[Fig 2: Leftmost and rightmost displays show stimuli to be remembered and matched for the memory test. Subjects search the central displays for a tilted line. In the uppermost central display, a non-target line appears inside a memory-matching red distractor shape. This stimulus draws attention away from the (target-) tilted line. (Soto et al. 2008)]

[12] Subjects' saccades may exhibit trajectories that are curiously bent toward the distractor, suggesting that it attracts attention, before its influence is suppressed midway by some other process. (Theeuwes et al. 1999; Walker & Sorley 2008; Theeuwes 2010)
[13] The same effect has been found for stimuli associated with rewards, such as food, which tend to attract subjects' attention without the subjects' knowledge. (Anderson 2013) The effect has been found for implicit memories of scene configurations, such as the co-occurrence of a target with a certain geometric arrangement in a scene, which will lead subjects to attend to locations in the arrangement that used to feature a target. (Chun 2003) And similarly, memories of scene-gist or typical locations for objects in scenes tend to make subjects attend to locations that do not, at present, contain the search target. (Torralba et al. 2006; Hollingworth 2014)

Even irrelevant items stored in working memory can interfere with subjects' search. (See Fig. 2) Suppose that we instruct subjects to maintain a red square-shape in working memory for later use in a memory test in which they determine whether a newly presented item matches the one they memorized. Suppose further that we interleave a visual search for a tilted line with the memory task, and that the search display features a red distractor matching the red square-shape held in working memory. The red item held in working memory can interfere with the subject's search. Search slows down, becomes less accurate, and features more errors. The item can interfere even when subjects know that it is a distractor. (Soto et al. 2005; Soto et al. 2008; Olivers et al. 2006) Nevertheless, individuals in these cases seem to guide their search: their search-goal drives; *they* direct where attention goes.

Generic appeal to the endogenous system thus does not give us a plausible candidate for constituting the individual's capacity to guide visual attention. In successfully guided visual search, different psychological sub-systems must be "brought in line," activated and regulated so as to contribute to, rather than interfere with, the individual's search. How is such regulation achieved? Empirical psychology appeals to regulation by the executive system to answer this question. (See Fig. 3) The executive system regulates assignments of priority on the priority map in light of the individual's goals. The system plays a twofold regulatory role. First, it helps set goals and provide goal-input to the computations determining priority on the map. Second, it enhances and inhibits influences on priority computations throughout the execution of the search, depending on whether they would increase or interfere with the likelihood of finding the search-target. (Eckstein 2011; Awh et al. 2012; Tsotsos & Kruijne 2014; Buschman & Miller 2014; Zelinsky 2015)



[Fig. 3: The executive system, as implemented by frontal and prefrontal cortex, accessing and regulating processing in other sub-systems for shifting attention. (Miller & Buschman 2013)]

Let me explain how the executive system regulates the influence of different sub-systems on priority-assignments in light of the birder's goal to find a robin. In light of her search-goal, executive switching helps set the search target, by encoding a representation of the target – a red, white, and brownish, thrush-shaped animal – into visual working memory. (Vickery & Jiang 2005) Executive maintenance holds this representation in working memory throughout the search, so as to provide top-down input to priority computations, enabling processes of template-matching. (Zelinsky 2008; Duncan & Humphreys 1989; Carlisle et al. 2011) Suppose that the birder also holds an irrelevant representation of a yellow warbler in visual working memory. Yellow objects in the scene should hence tend to attract her attention, interfere with her search. To counteract such interference, executive allocation of processing resources may boost the activation of the search target in memory, thereby increasing its influence on priority computations. (Olivers & Eimer 2011) Alternatively, executive inhibition may suppress the influence of an interfering factor, such as a flower whose yellow color matches that of the irrelevant working memory item. (Sawaki & Luck 2011) In either way, the executive system can regulate the influence of endogenous factors so as to avoid their interference with the individual's guidance. But suppose that the robin is also the only red object in some corner of the scene. The exogenous system boosts its location on the priority map due to its saliency. Since the robin's salient feature is congruent with the search-goal, in light of which the executive system regulates, the latter does not inhibit its influence on the priority map. In this way the executive system can recruit the exogenous system to benefit the individual's guidance. (Folk et al. 1992) Executive regulation takes time and relies on limited resources to establish its regulatory influence on priority computations. When the individual has insufficient time, if interfering factors override executive regulation, or if some particularly salient stimulus shuts down the influence of executive regulation of the priority map, we obtain the phenomenon of captured attention. (Han & Kim 2009)

Let me provide some evidence for these claims about executive regulation.

Executive maintenance in working memory of information about the search target provides top-down information about search-goals to priority computations. We know from computational models of visual search that a target-template is held in memory, throughout the execution of a search, so as to enable a process of template-matching between search-target and items in the scene. (Najemnik & Geisler 2009; Eckstein 2011; Zelinsky 2008, 2015) Why think that the relevant memory is an executive working memory? Evidence comes, for instance, from a range of studies of brain activity during search. Thus Carlisle et al. 2011 measured event-related potentials (ERPs) in subjects performing visual searches. A component of this wave, the contralateral delay activity (CDA), has been established as indicating active maintenance of object representations in visual working memory. (Vogel & Machizawa 2004; Vogel et al. 2005) Subjects in the study by Carlisle et al. first saw a cue-display that indicated the shape of the target object. The researchers measured ERP-waveforms following cue-presentation. Carlisle et al. found the CDA-component in the ERP wave. These findings provide strong evidence for the stipulated role of visual working memory in maintaining templates for visual search. (Oh & Kim 2004; Woodman & Luck 2004; Woodman & Arita 2011; Beck 2012; Dube et al. 2016)

Executive switching helps establish and change the set for a search. Vickery & Jiang showed their subjects a search cue for 200-500 ms before the subjects searched a display. The cues could be 2D-, 3D-images, or words describing the search-target. Vickery & Jiang varied the time subjects had to set up the search-target between 200-

1000 ms. Vickery & Jiang found that setting up an exact visual search target in working memory took around 200-500 ms while setting up a target on the basis of a verbal cue took up to 1000 ms. Both results are consistent with research on the cost induced by executive switching generally. They thus support a role of executive switching in visual search. (Vickery & Jiang 2005; Walther & Li 2007; Dombrowe et al. 2011)

Executive allocation of processing resources helps increase the influence of task-relevant factors over distractors. Suppose that subjects maintain several items in working memory, some relevant to an ongoing search, others not. Olivers & Eimer provided evidence that task-relevant representations in working memory may be preferentially activated or boosted so as to decrease working memory interference. Olivers & Eimer asked their subjects to first encode a color – say, red – into working memory for a subsequent memory test. Next their subjects performed a visual search for a tilted line among distractors.[14] One of the distractors would match the color (red!) maintained in working memory. We have seen earlier that under these conditions, working memory can draw attention. Olivers & Eimer manipulated the order in which working memory test and visual search would be performed. When both tasks were randomly intermixed, they reasoned, both task sets should be kept more active than when subjects can predict which task would be performed next. The effect of the memory distractor on visual search should be greater in the intermixed condition. This prediction was confirmed. Reaction time costs due to memory-matching distractors more than doubled (from 30 ms to 65 ms) and subjects made 1.5% more errors. (Olivers & Eimer 2011; Van Moorselaer et al. 2014, exp. 2 & 4; Han & Kim 2004)

Executive inhibition can suppress salient stimuli or distractors matching irrelevant working memory items. Sawaki & Luck 2011, too, investigated components of the ERP waveform during visual search. Just as the CDA-component indicates visual working memory activity, the N2pc-component is associated with the allocation of attentional resources to a stimulus. A slightly less familiar component has been shown to reflect an inhibitory mechanism – the Pd-component. (Hickey et al. 2009; Sawaki & Luck 2010) Sawaki & Luck presented their subjects with two colored objects. A spatial cue indicated which object to maintain in memory for a subsequent memory task. While the cued

[14] Displays resembled those in Fig. 2.

object was maintained in working memory, two task-irrelevant circles were briefly presented. One matched the color of the item in working memory, the other did not. Sawaki & Luck documented the Pd-component for the memory-matching distractor. The finding suggests that the distractor exerted an attraction on visual attention that was inhibited before attention could shift to the distractor. Subsequent studies found this component for distractors during visual search. This research provides evidence for the role of executive inhibition in visual search. (Sawaki & Luck 2013; Mertes 2016; cf. also Emrich et al. 2010; Han 2015; Dube et al. 2016; Lu et al. 2017)

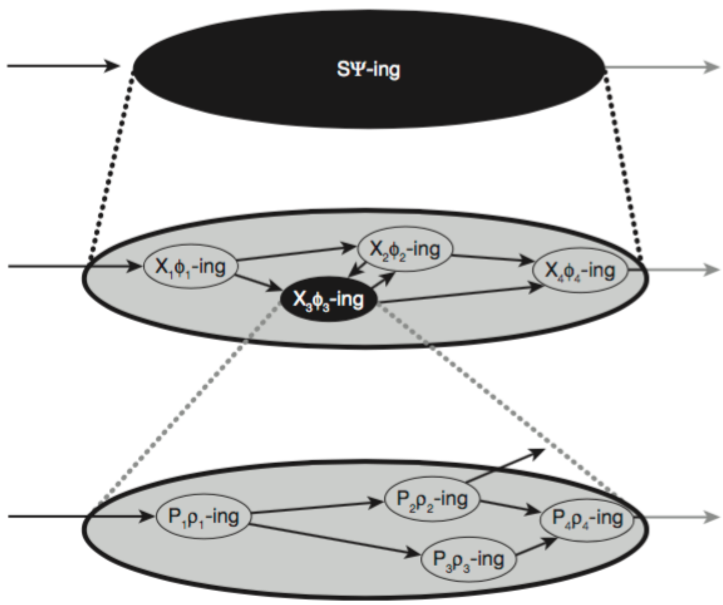
A wide range of sub-systems thus helps orient visual attention. The executive system regulates these sub-systems' influence on attention-shifts, in light of individuals' goals. Such regulation is absent from passive, non-guided shifts. Executive regulation thus correlates with individuals' guidance of visual attention at the level of psychological sub-systems. Its role in orienting attention makes it a plausible candidate for constituting a capacity to guide.

5 Capacity-constitution

Suppose that executive regulation indeed correlates with exercises of individuals' guidance. One may still wonder *how* this fact bears on individuals' guidance. My claim is that the executive system actually constitutes individuals' capacity to guide visual attention. What does this claim of capacity-constitution amount to? How can we establish that executive regulation is not merely a causal enabling condition on, but actually *constitutes* a capacity to guide visual attention? How do these claims *explain* guidance? In this section I focus on answering these three questions. In doing so I rely on ideas from recent philosophy of cognitive science.

Let me address the first question first: What does this claim of capacity-constitution amount to? To *actually constitute* a capacity to guide visual attention, the executive functions must be components of that capacity. To be components of the capacity, the executive functions must figure in a componential explanation of individuals' capacity to guide. The ability to *explain* a target-phenomenon is thus built into conditions on capacity-constitution.

An explanation is *componential* if it decomposes a target phenomenon into its components, and explains the target phenomenon as generated by these components' organized causal interaction.[15] (See Fig. 4) It aims to explain how some target phenomenon can arise, given the causal structure of the actual world. Thus I claim that the executive functions are components of individuals' actual capacity to guide their activities – their activity of shifting attention in particular – toward their goals. Componential explanation really is a family of views that comprises what has been called functional explanation, mechanistic explanation, or explanation by cognitive models. (Cummins 1983; Machamer et al. 2000; Craver 2007; Bechtel 2008; Weiskopf 2011 & 2018)[16]



[Fig. 4: Illustration of how processes at different explanatory levels help constitute an individual-level capacity. (Craver 2007, 189)]

[15] This characterization relies on Craver 2007, esp. chapter 4, and Weiskopf 2018.

[16] There are important differences between members of this family. One dispute between family-members concerns whether all componential explanation must (eventually) explain in terms of neural mechanisms. Some views carry a strong commitment to this effect, others do not. (Craver 2007; Piccinini & Craver 2011; Weiskopf 2018) While I am sympathetic to the latter position, I do not think that anything I say in what follows depends on it.

This use of 'constitutes,' while standard in philosophy of cognitive science, differs from some uses in the wider literature.[17] Claims about actual constitution differ, for instance, from those concerning metaphysical constitution. I do not make any modal claims. In particular, I do not claim that executive regulation is a necessary condition on individuals' guidance, across organisms and possible worlds. Other organisms in the actual world, as much as agents in other possible world, may act, and guide their activities, but not have an executive system. Claims about actual constitution also differ from claims concerning supervenience, implementation, or realization. (Kim 1998) Claims about actual constitution are both weaker and stronger. They are weaker in that they do not purport to provide reductive sufficient conditions on guidance, even in the nomological sense.[18] The conditions are stronger in that they maintain that actually constitutive conditions *explain* individuals' guidance, given the causal structure of the actual world. There is no standard constraint on supervenience, implementation, or realization, that the respective conditions explain a target-phenomenon.

With this rough characterization of a notion of capacity-constitution in hand, we can now turn to addressing the second question. How can we *establish* that executive regulation actually constitutes individuals' capacity to guide visual attention? There are surely different ways to do so. On one influential way of arguing this, to figure in componential explanation, or to be real components of a capacity, components must be robust, and they must play an actual causal role in generating the target phenomenon. (Craver 2007, 4.5 & 4.8; Weiskopf 2018) The executive functions meet both criteria. I here rely on the empirical facts laid out in section 4.

Real components are *robust*: they have a stable cluster of properties, and are detectable in a variety of causally and theoretically independent ways. (Craver 2007, 132)[19] Are the executive functions robust in this sense? They are.

Executive switching exhibits characteristic time-courses and error-patterns. Switching from one task-set to another takes time. How much time switching takes

---

[17] Kim 1998; Burge 2010
[18] Executive regulation is at most a core realizer of individuals' guidance that must be embedded in a range of other conditions. Only a total realizer is sufficient for some phenomenon. (Shoemaker 1981)
[19] Craver mentions intervention and physiological plausibility as further requirements. Intervention is connected to mutual manipulability. Physiology either is non-essential or trivial as a requirement. (Weiskopf 2016) I hence focus on the two features mentioned in the text.

depends on the task-set involved. Establishing a visual target-template, as we saw, takes between 200-500 ms. For verbal templates, switching takes about 1 s. (Vickery & Jiang 2005) Error-patterns can be distinguished as due to preparatory, response, or stimulus-interference. (Kiesel et al. 2010) Switching especially involves dorsolateral prefrontal cortex, anterior cingulate, and posterior parietal cortex. (Hakun & Ravizza 2008; Ravizza & Carter 2012; Li et al. 2012)

Executive working memory maintenance exhibits fairly stable capacity limits of around four items. (Vogel et al. 2001) It exhibits characteristic time-courses for encoding information (20 – 100 ms per item). (Bays et al. 2011) Executive maintenance fractionates into stores for visual, spatial, action-related, and verbal information, stores that operate on different types of representations. (Baddeley 2007) Modulation of activity in different brain areas by dorsolateral prefrontal cortex likely implements working memory maintenance. (D'Esposito & Postle 2015)

Executive inhibition, too, exhibits capacity limits. We find characteristic breakdown-patterns of inhibition depending on strength of distractor-salience, and perceptual or executive load induced in dual-task paradigms. (Lavie & Dalton 2014; Han 2015) Inhibition also exhibits characteristic time-courses of operation. Thus executive inhibition takes about 500-750 ms to fully establish its control over distractors or prepotent responses across paradigms. (Han & Kim 2009) The inhibition function seems to be implemented primarily by right inferior frontal cortex-modulation of other areas in the brain. (Aron 2004; Munakata et al. 2011)

Executive resource-allocation similarly exhibits capacity limits. Concurrent executive processing interferes with primary tasks relying on executive resource-allocation. (Diamond 2013) Whether we find interference depends on the kind of task – mathematical, verbal, visuo-spatial – performed. (Baddeley 2007) Resource-allocation again is a matter of modulating neuronal activity in different brain areas. Resource-allocation to working memory representations, for instance, apparently involves the endogenous increase in firing of neurons in dorsolateral prefrontal cortex. (Olivers et al. 2011; Warden & Miller 2007, 2010)

Each of these competencies is furthermore detectable in a variety of independent ways: a vast range of behavioral paradigms is used to study the executive functions, including delay-period, n-back paradigms, dual task paradigms, Stroop, Simon, or Antisaccade tasks, as well as mixed, single task, predictable, and unpredictable switching tasks, and so on. (Baddeley 2007; Kiesel 2010; D'Esposito & Postle 2015; Diamond 2013) Computational models study and yield predictions about characteristics of each of the four competencies. (Botvinick & Cohen 2014) A wide range of neuroscientific methods such as single neuron-recording, fMRI, TMS, and ERP-studies helps identify the competencies' operation. (Miller & Buschman 2013; Li 2012; D'Esposito & Postle 2015; Aron 2004; Warden & Miller 2010)

So the executive functions are robust, or meet criteria for being real components of a capacity to guide.

But do the executive functions play an actual causal role in individuals' guidance of their activities toward their goals? Yes, they do. We can establish this by showing that individuals' guidance and executive regulation are mutually manipulable. (Craver 2007, 139ff.; Campbell 2008; Weiskopf 2018)[20]

> *Mutual manipulability*: the target component is a part of the capacity and (i) interventions on exercises of the target capacity change activities of its components, and (ii) interventions[21] on the activities of components change exercises of the target capacity.

Executive regulation and individuals' guidance *are* mutually manipulable in this sense. To appreciate this, we must remember the evidence concerning the correlation of individuals' guidance and the exercise of different executive functions, discussed in the last section.[22] Let me indicate in slightly more detail how the argument for each executive function would go.

First, consider executive switching. When an individual guides visual attention during visual search toward finding a green diamond, the search typically requires switching from a prior task to the present one. Initiating the search thus will involve

[20] For critical examination of this idea see Baumgartner 2017.
[21] Craver 2007, 96: "An ideal intervention on X with respect to Y is a change in the value of X that changes Y, if at all, only via the change in X. More specifically, this requirement implies that: (i) I does change Y directly; (ii) I does not change the value of some causal intermediate S between X and Y except by changing the value of X; (iii) I is not correlated with some other variable M that is a cause of Y; and (iv) I acts as a 'switch' that controls the value of X irrespective of X's other causes, U."
[22] See especially pp. 16-9.

executive switching. An empirical signature of switching is a component in the ERP-curve: the cue-locked switching positivity in posterior areas about 300 ms after cue onset, as well as a stimulus-locked switch negativity. Both are evidenced in the ERP-wave in standard visual search paradigms. (Li et al. 2012; Nicholson et al. 2006) Thus engaging in guided visual search activates the component-competence of executive switching.

Interfering with executive switching, on the other hand, will yield interference with the individual's guidance. As mentioned earlier, setting up a visual target requires about 200-500 ms. Giving individuals insufficient time to initiate a visual search will interfere with their competence for switching. This interference will lead to an insufficiently consolidated search-set, causing in turn an increase in errors and reaction time. (Vickery & Jiang 2005) Interference with executive switching thus yields interference with the individual's goal-directed guidance.

Next, consider executive maintenance. As we have seen, if an individual guides visual attention during visual search toward finding a green diamond, then a representation of the green diamond-shape will be stored in working memory. ERP studies show CDA-activity for such search paradigms. (Carlisle et al. 2011) So activating the target capacity activates the component – working memory, in this case.

But similarly, intervening on working memory load affects the efficiency with which an individual guides her visual attention toward finding a green diamond target. Thus there is evidence that load on spatial working memory or visual working memory increases reaction times and the extent to which distractors draw attention. (Han & Kim 2004) So interference with working memory activity interferes with the individual's goal-directed guidance.

The same is true for executive inhibition. If an individual guides visual attention in search for a green diamond shape, then her guidance will involve inhibition of distractors in the display. Evidence from ERP-studies shows that such search activates a Pd-component in the ERP-wave, which implements the inhibition of distractors. (Sawaki & Luck 2011) So the individual's engagement in goal-directed guidance of attention in visual search will tend to activate executive inhibition.

And again, interfering with the inhibitory component yields interference with the individual's guidance of her search. We know that some inhibitory effects depend on information held in working memory. Loading working memory yields interference with inhibition of distractors. And such interference negatively affects individuals' guidance of their search: reaction times increase, errors increase, and individuals orient attention more often to irrelevant distractors. (Beck et al. 2012; Lu et al. 2017) So interfering with executive inhibition we here interfere with individuals' goal-directed guidance.[23]

Finally, consider executive resource-allocation. We saw that, when individuals guide their visual attention in search for a green diamond shape, then the activation of the search target's working memory representation may be boosted. Additional executive resources are then allocated to the representation of the search target. (Olivers 2010) The individual's engagement in goal-directed guidance of attention in visual search activates executive resource-allocation.

But on the other hand, interfering with executive resource-allocation yields interference with the individual's guidance of her search. Many studies have shown that, if individuals carry out a visual search while concurrently performing a task that requires executive resources – such as counting backwards – then this secondary task will interfere with the individual's guidance. Her search becomes slower and she is prone to more errors. (Han & Kim 2004) So interfering with executive resource-allocation tends to yield interference with individuals' goal-directed guidance.

So the executive functions do play an actual causal role in individuals' guidance of visual attention. If the argument so far has been correct, we have thus directly established that the executive functions actually constitute an individual's capacity to guide. This, in turn, allows us to address the last of our three questions: How do claims about constitution explain some target-phenomenon? Insofar as capacity-constitution, on this conception, has claims about explanation built into it. Components of a capacity are elements in componential explanations of that capacity. The causal interaction of capacity-components constitutes exercises of the whole individual's capacity. We thus

[23] A reviewer suggests that my argument here might be problematically relying on reverse inference. (Poldrack 2006) It does not. My argument takes at face value empirical results to the effect that ERP-signatures indicate working memory involvement in specific task-contexts. It then uses these results in arguing for mutual manipulability. Does the empirical work itself rely on reverse inference in a problematic way? I see no reason for thinking that. The researchers rely on a wide range of complementary methods and converging evidence to argue for the ERP-signature as an indicator of working memory, over alternative psychological processes.

explain how individuals guide their visual attention-shifts by explaining how the causal interaction of the executive functions between them, and their regulation of other sub-systems, generates individuals' guidance.

6 Some worries

Let me address some worries that may naturally arise at this point. The first worry concerns whether my proposal over- or under-generates. The second charges that my explanation of guidance is circular. The third revives concerns about disappearing agents. The fourth claims that I am missing the explanatory target altogether.

Over-/under-generation. Does the proposal over-generate? One version of the worry might take it that any exercise of executive functions generates an episode of guidance. Does any encoding of information into working memory, for instance, constitute an agent's act? Surely, that would over-generate guidance. But the view does not commit to his idea. The view focuses on guided visual attention-shifts. For these exercises of agency, the specific causal cooperation of different executive functions and their interactions with other sub-systems – their top-down regulation through switching, working memory, inhibition, and allocation – are required for the individual to guide.

Another version of the worry charges that the criterion of mutual manipulability over-generates. Would we not interfere with individuals' guidance if we interfered with individuals' visual processing, by damaging V1? If we damaged V1, visual information could not serve as input to computations of priority. In many cases, lack of this information would affect individuals' search. Does processing in V1 hence help constitute a capacity to guide visual attention? No. Visual processing in V1 surely is a factor in many explanations of how individuals shift attention. But such damage is too unspecific to indicate a component that helps constitute a capacity to guide (Craver 2007, 157ff.): it affects not merely the individual's capacity to guide attention, but many passive visual perceptual episodes as well. Similarly, while it will interfere with individuals' search in many cases, it need not affect their search in all relevant cases. For suppose that the individual *remembers* a search target's likely location. The individual can then shift her attention to that location without relying on visual input for her shift.

This type of over-generation worry can be allayed by pointing to the empirical evidence (see section 4) that only the relevant kind of executive regulation correlates with individuals' guidance of visual attention. To the extent that this claim is accurate, only such regulation provides a plausible candidate for constituting this capacity to guide. While executive regulation relies on and interacts with a wide range of different sub-systems – after all, it is the whole agent that must act – only the executive system operates across all the guided episodes.

Does the proposal under-generate guidance? Don't individuals, for instance, guide their motor actions by relying on the operation of their motor system? How about systems for adaptive motor learning? While I do believe that executive regulation constitutes individuals' guidance across *all* the different ways in which they can act, I have not argued that claim here. Here, I focus on establishing the claim that executive regulation constitutes a capacity to *guide shifts of visual attention.* The claim's truth is consistent with the existence of *other* capacities to guide – for instance, capacities to guide motor behavior – constituted by other sub-systems.

Circularity. One might worry that explanations of individuals' guidance in terms of the executive system are, in effect, circular. Aren't we, or the scientists, just assuming that the executive system is "whatever system explains actual human agency"? Or alternatively, doesn't circularity enter the picture because we, in effect, re-introduce an agent with agential powers?

There is no assumption that the executive system is "whatever system explains agency." Indeed, the science debates about the proper function of the executive system. Some theorists identify its function as that of enabling cognition, or thought. (Baddeley 2007) Others associate it with goal-directed behavior. (Gazzaniga 2019; Badre 2020) Neither assumption plays a role in the science's actual explanations of specific psychological episodes. These episodes are explained in terms of the interaction of the executive system's component-competencies – the executive functions. Their explanatory status is autonomous, because they are robust. They are independently characterized, have a stable cluster of properties, and are detectable in a variety of independent ways. As we have seen, the claim that the executive system constitutes a capacity to guide requires extensive argument – from correlation to constitution. The first circularity-worry is misdirected.

But aren't the executive functions endowed with powers so great that I, in effect, re-introduce the agent and her agential powers? First of all, the executive functions are *components* of a capacity to guide. The latter is an individual's agential capacity. The former are parts of a psychological sub-system that helps constitute this capacity. Individual executive functions are not plausibly, certainly not obviously, agential capacities, rather than competencies at the sub-system level. But second, even if they were, they are not, individually, capacities to guide. To the extent that my argument about capacity-constitution is correct, the executive functions must interact, in coordinated ways, with one another, for the individual to guide. So even if, individually, they were agential capacities, we could still appeal to them in explaining individuals' guidance. Finally, I argued that the executive functions together constitute *one* of an individual's agential capacities. Even so do they not constitute the whole agent. For a whole agent to act, her many capacities, agential and otherwise, need to interact. So there is no threat of vicious circularity.

Disappearing agents. Does the proposal not revive worries about disappearing agents? It seems as though I propose to explain individuals' guidance in terms of the operation of some specific sub-system – the executive system. But sub-systems are not agents, as per my own argument above. So have we not, if the arguments are successful, abolished a role for the agent in action?

The same considerations that helped us address circularity-worries can also help us clear up this misunderstanding. First, I do not claim to have provided a reductive account of individuals' capacity to guide, in terms of executive regulation. Not only is it, at this point, unclear whether this whole individual-capacity *can* be reduced to interactions of the executive functions: this may not be possible in principle (Craver 2007, 196ff.); we certainly do not have such a reduction as a matter of fact, given how incomplete our understanding of this system is. Second, even if such a reductive account of individuals' capacity to guide attention-shifts were available, I would not have explained away the agent. For the whole agent to shift her attention, her capacity to guide must interact with other capacities, such as those for memory, visual perception, and so on. For the whole agent to be, other capacities and sub-systems also must be in place. They must interact for whole-individual action to occur. So appeals to this system as constituting a capacity to guide do not make the acting individual disappear.

Wrong target. One might finally worry that the proposal entirely misses the explanatory target of traditional attempts to explicate agency. These attempts aim for agency's essence or nature. They are not interested in the actual constitution of human agency.

I conceded that the present inquiry does not give us that. But this does not mean that it is without interest: to us, to action theory, or even for the quest for agency's nature. First, understanding actual human agency is a topic worthy of our interest, and certainly within action theory's province. But second, traditional theorizing about agency has not been able to resolve issues such as the problem of guidance, disappearing agents, and causal deviance. Understanding these phenomena in the actual world may provide us a novel take on old issues, one that can direct philosophical reflection, even about agency's nature. I thus propose to see these considerations as widening the philosophical repertoire. We should explore rather than dismiss them.

7 Two recent proposals[24]

Recent years have seen proposals by other authors as to the difference between actively guided and passive attention-shifts. Let me briefly comment on two. The first proposal is by Zachary Irving:

> "[S]omeone's attention is directed (that is, guided) by a goal only when she would feel pulled back from distractions. More precisely, an agent A's attention is directed by one of her goals, t, if and only if A is guided to focus her attention on information that she takes to be relevant to t; that is, A has two dispositions:
>
> 1. A reliably focuses her attention on information that she takes to be relevant to t and
> 2. If A's attention weren't focused on information that she takes to be relevant to t, A would feel distracted and thereby be disposed to correct this fact." (Irving 2021, 622)

[24] A reviewer insists that I compare my position with these two proposals. I thank them for providing a pdf of Irving's paper at a time when it was still unpublished. References are to the published version. Space does not permit a fuller discussion of these proposals.

Let me begin by briefly stating how I see my position in relation to Irving's. I would want to resist several aspects of his account. I think that individuals can engage in actively guided visual search while being extremely unreliable at focusing on relevant information, or information they take to be relevant. The birder from section 2 may have her attention drawn to leaves, branches, and so forth, none of which are (or she takes to be) relevant to her goal of finding the robin. I furthermore think that she can engage in actively guided visual search without representing information as (ir)relevant to her search. That is, I believe that she can guide without "taking" information to be relevant to t. Finally, I would want to resist a commitment that she must feel any particular way, when fixating irrelevant information.[25] The executive system might regulate the birder's search, absent the relevant feeling. To be sure, this is merely to put the disagreement into relief. I here do not have the space to appropriately argue these points.

Instead I focus on Irving's criticism of a view similar to the present one. He writes:

> "Buehler [2014] holds that executive control is constitutive of guided attention, in the actual world and for creatures like us. In contrast, I argue that cases like skilled guided attention bypass executive control. We must look to action theory, not neuroscience, to see what unifies guided attention in "creatures like us."" (Irving 2021, 636)

Although my topic in the present paper was not skilled action, I consider the case of visual search discussed earlier an exercise of skilled, actively guided attention. So I take the argument in this paper to support the claim that executive regulation is operative in the active guidance of skilled action. Several authors have recently argued for this position at length. (Christensen et al. 2016; Fridland 2021; Buehler 2021a) I won't repeat these arguments here, as this would lead us too far afield.

Why does Irving object to this position? I am unsure: his own proposal seems consistent with skilled, actively guided attention's being regulated by the executive system. While Irving does not explain what he takes skill to be, he gives the example of a

[25] Irving adopts this requirement to account for guidance's being individual-level. (Irving 2021, 631) I have argued that we can explicate individuals' guidance by appeal to executive regulation, which need not involve the feelings at issue here. (Cf. sections 5 & 6, p. 28, of this paper, as well as Buehler 2021b, section 5.4)

"skilled fisher, who guides her attention toward signs of fish (for example, reeds and ripples. […]" (Irving 2021, 632; see also pp. 629 & 635) If the fisher actively guides visual attention, then this is precisely a case of visual search, regulated by the executive system.

In the same context, Irving also discusses empirical work[26] that he says shows that "constraints from control *or* salience networks both result in guided attention." He claims that certain

> "[c]ases of guided attention bypass control networks and instead result from the salience network's constraints on the DNMTL. The salience network may support skilled actions, where we effortlessly guide attention to salient, goal-relevant information. When a skilled fisher effortlessly imagines where to anchor her boat, for example, her salience network may constrain her DNMTL. As a result, she may effortlessly guide her attention to anchoring spots." (Irving 2021, 635)

Is the idea that the empirical work shows that in actively guided, skilled attention shifts, a salience-network *alone* orients attention, and therefore, the executive system does not regulate? The empirical work cited by Irving does not support this idea. None of the articles that Irving references even mention "skill." The empirical work argues that control- and salience-networks are involved in mind-wandering to a lesser degree than in other cases, while the DNMTL is involved to a greater degree. (Christoff et al. 2016, 7, 9 Fig. 5. & 11; Seeley et al. 2007, 7) Nothing I say commits me to rejecting this claim. Nor do I reject the idea that the "salience network … support[s] skilled actions." I explicitly acknowledge a role for saliency in actively guided, skilled attention-shifts. (See section 4.) I merely insist that in actively guided skilled attention, saliency *interacts* with executive system regulation. As far as I can see, the empirical work is compatible with my claims concerning executive regulation in actively guided attention-shifts.

A reviewer suggests that the idea might be rather that guidance can be passive or active, and that cases of *passively* guided attention can "bypass control networks." If so, then what Irving writes is compatible with my proposal, which concerns actively guided attention-shifts only.

---

[26] Especially Christoff et al. 2016. See also Irving 2016; Dixon et al. 2014; Hanna et al. 2018; Irving & Glasser (2019); Murray, Irving & Krasich (forthcoming)

The second proposal is by Sebastian Watzl. Irving appears to appeal to Watzl as "agreeing" with him (as against the view in Buehler (2014)) that "a disjunctive set of neural mechanisms result in guided attention, including executive control and salience." (Irving 2021, 636) Does Irving mean that the *active* guidance of attention-shifts can be realized by a disjunctive set of neural mechanisms, including either the saliency, or the executive control-network? If so, then it seems that Watzl disagrees. Watzl indeed describes cases of attentional capture, for instance, as "passively guided attention." (Watzl 2018, 117) What here guides the attention-shift, for Watzl, is an individual-level perceptual (or other) state of the individual. He calls this the

> "*Subject-level guidance claim*. When a subject's attention is passively guided it is guided by subject-level states." (Watzl 2018, 115)

But Watzl distinguishes such passive guidance from the active, goal-directed guidance by the individual at issue in the present article. He writes:

> "When it is guided by psychological salience, attention is not controlled by the subject. In this way, attention that is guided by psychological salience contrasts with actively guided attention." (Watzl 2018, 114)

He continues:

> "*Active guidance*. What it is for a subject's attention to be actively guided is for her attention to be guided by her executive control system." (Watzl 2018, 140)

Watzl describes captured attention as passively *guided*, because he wants to emphasize that attend*ings* – episodes of paying attention to stimuli (even after they passively captured a shift) – have a special metaphysical status: they are activities. (Watzl 2018, 66ff.) This claim is consistent with my proposal. I do not make claims about episodes of attending, but rather focus on actively guided attention-shifts. As far as I can see, Watzl's position is not only compatible with my view concerning executive regulation in actively guided attention-shifts, but explicitly builds on it.[27]

[27] See his Chapter 7, Fns. 9, 29, 30. I like to think of Watzl 2018 as a way of integrating ideas on executive regulation and priority maps from Buehler 2014 with the theory of attention as structuring consciousness, proposed in Watzl 2010.

Or is the idea rather that guidance can be passive or active, and that Watzl agrees that these different forms of guidance are respectively realized by different neural mechanisms?[28] That seems to indeed be Watzl's view. But on this reading, the passage is compatible with the position I defended in the present paper, too: I have argued that *actively* guided attention-shifts are executive-regulated. I do not comment on passively guided attention-shifts.

8 Conclusion: explaining individuals' goal-directed guidance

In this paper I have attempted to make progress on the task of explicating agency, by explaining individuals' guidance. I have argued that actual human agents have a primitive capacity to guide their attention toward a goal; that this capacity is constituted by an empirically discovered sub-system – the executive system; and that we can explain guidance, by investigating the capacity's sub-individual constitution.

One aim was to explore how empirical research might help us better understand individuals' direction of their actions toward some goal. The methodology faces two challenges. It must first find the elements in empirical explanations that are plausible candidates for explaining individuals' guidance. Next, it must argue that the research indeed bears on philosophical questions concerning such guidance. I showed that reflection on the role of different psychological sub-systems in explaining how individuals shift visual attention allows us to identify the processes that are likely candidates for bearing on our philosophical question. Next, I showed that we can use considerations concerning mutual manipulability to *directly establish* that some candidate processes bear on philosophical questions about agency: they actually constitute the individual's capacity to guide. This strategy allows us to establish that some proposed condition – executive regulation – helps constitute the *individual's* guidance. The methodological upshot is twofold. We can indeed make progress on philosophical questions about agency, and especially concerning how *individuals* guide, by investigating the sub-individual constitution of agential capacities. We make progress by

[28] Thanks to the reviewer for suggesting this reading.

immersing ourselves in case studies and by providing an empirically-based philosophical argument for capacity-constitution.

What have we learned about the capacity to actively guide visual attention? In section 2 I sketched surface-properties, or marks of guidance. They served to circumscribe our target-explanandum. This raised a host of further questions: Is there *really* such a primitive capacity? How, and on what grounds, can we further characterize this capacity? How can we explain its operation in an illuminating, non-circular way? If my argument has been successful, then all these questions have found answers. Sections 4 and 5 provided grounds for thinking that such a capacity to guide does exist. It is the capacity constituted by the executive system. Empirical investigation into this system and its operations provide ways to characterize the capacity: the capacity is constituted by the component-capacities of switching, maintenance, inhibition, and allocation. Further empirical investigation into these component-capacities will yield further progress in understanding them, and through them, the explanatory target – our capacity to guide visual attention. We explain the operation of this capacity precisely by explaining the operation of the system in orienting attention – the causal interaction of the different executive functions, and of the executive system as a whole with other sub-systems. This explanation is illuminating, and indeed, non-circular, because the executive functions, as much as other sub-systems, are robust, independent explanatory posits, that provide grounds for a flourishing science.

What progress, then, have we made explicating agency? The task was to identify a causal mechanism whose operation would (help) constitute the agent's active guidance – as she regulates her activities, *throughout* the execution of her action. When the individual guides, Frankfurt said, her

> "behavior is … under the guidance of an independent causal mechanism, whose readiness to bring about compensatory adjustments tends to ensure that the behavior is accomplished. The activity of such a mechanism is normally not, of course, guided by us. Rather it is, when we are performing an action, our guidance of our behavior." (Frankfurt 1978, 74/5)

Where Frankfurt merely hypothesized that there should be such a mechanism, we have directly identified, and described such a mechanism. Traditional proposals were open to the charge launched from the beginning of this paper, of justifying *why* some causal mechanism should constitute an agent's guidance. The response is that this is a brute empirical fact. We have, in effect through ostension, picked out the relevant mechanism. We know that it constitutes the *agent's* guidance, because we can directly test this hypothesis. We might now attempt to marshal these results to address challenges from disappearing agents or deviance.

When the individual actively guides her behavior, then the executive system regulates processing across her sub-systems, throughout the execution of the action. When the individual does not so guide, then the executive does not so regulate. We thus have, at least for visual attention shifts, an informative way to distinguish active from passive episodes. I believe that this proposal can be made to work for other kinds of action, too. It is easy to appreciate this idea's initial plausibility: across act kinds, activity across sub-systems has to be coordinated, integrated, interference has to be suppressed, so as to ensure direction of the activity toward some goal. Regulation by the executive system seems to be of the structurally right kind to achieve this kind of whole-individual coordination. By holding the action's goal in working memory, initiating processing across sub-systems for attaining the goal through switching, and regulating processing across sub-systems throughout the action, so as to ensure goal-attainment, the relevant whole-individual coordination could be realized – whether the relevant action involves moving one's arms, one's legs, or solving a problem in arithmetic. I hope to develop this idea on another occasion. Here, my aim was to argue for a capacity to guide visual attention, and to validate this kind of approach for action theory.

Author:
Denis Buehler
Institut Jean Nicod

École normale supérieure, Paris

France

Literature

Adams, F. & Mele, A. (1989) “The role of intention in intentional action”, *Canadian journal of philosophy* 19(4): 511-31.

Aron, A., Robbins, T. & Poldrack, R. (2004) “Inhibition and the right inferior frontal cortex”, *Trends in cognitive sciences* 8(4): 170-177.

Allport, A. (1993) “Attention and control: have we been asking the wrong questions?”, in Meyers & Kornblum (eds.), *Attention and performance XIV (pp. 183-216)*, Cambridge, MA: MIT Press.

Anderson, B. (2013) “A Value-Driven Mechanism of Attentional Selection”, *Journal of Vision* 13(3): 1-16.

Awh, E., Belopolsky, A. & Theeuwes, J. (2012) “Top-Down Versus Bottom-Up Attentional Control: A Failed Theoretical Dichotomy”, *Trends in Cognitive Sciences* 16(8): 437-43.

Baddeley, A. D. (1986) *Working Memory* New York: Oxford University Press.

Baddeley, A. D. (2007) *Working memory, thought, and action* Oxford: Oxford University Press.

Badre, D. (2020) *On task* Princeton: PUP

Bays, P., Gorgoraptis, N., Wee, N., Marshall, L. & Husain, M. (2011) “Temporal dynamics of encoding, storage, and reallocation of visual working memory”, *Journal of vision* 11: 1-15.

Bechtel, W. (2008) *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience* London: Routledge.

Beck, V., Hollingworth, A. & Luck, S. (2012) “Simultaneous control of attention by multiple working memory representations”, *Psychological science* 23(8): 887-898.

Bishop, J. (1989) *Natural Agency. An Essay on the Causal Theory of Action* Cambridge: Cambridge University Press.

Botvinick, M. & Cohen, J. (2014) “The computational and neural basis of cognitive control: charted territory and new frontiers”, *Cognitive science* 38: 1249-1285.

Brand, M. (1984) *Intending and acting* Cambridge, MA: MIT.

Bratman, M. (1987) *Intentions, plans, and practical reason* Cambridge, MA: Harvard.

Bratman, M. (2007) *Structures of agency* New York: OUP.

Brent, M. (2017) "Agent causation as a solution to the problem of action", *Canadian journal of philosophy* 47(5): 656-673.

Bruce, N. & Tsotsos (2009) "Saliency, attention, and visual search: an information theoretic approach", *Journal of vision* 9(3): 1-24.

Buehler, D. (2014) *Psychological agency – guidance of visual attention* UCLA Dissertation.

Buehler, D. (2017) "The central executive system", *Synthese* 195(5): 1969-91.

Burge, T. (2009) "Primitive agency and natural norms", *Philosophy and phenomenological research* 79: 251-278.

Burge, T. (2010) *Origins of objectivity* New York: Oxford University Press.

Buschman, T. & Kastner, S. (2015) "From behavior to neural dynamics: an integrated theory of attention", *Neuron* 88: 127-144.

Buschmann, T. & Miller, E. (2014) "Goal-direction and top-down control", *Philosophical transactions of the royal society* 369: 1-9.

Campbell, J. (2008) "Interventionism, control variables, and causation in the qualitative world", *Philosophical issues* 18: 426-445.

Carlisle, N., Arita, J. Pardo, D. & Woodman, G. (2011) "Attentional templates in visual working memory", *Journal of neuroscience* 31: 9315-9322.

Carnap, R. (1950) *Logical foundations of probability* Chicago: University of Chicago Press.

Christensen, W., Sutton, J. & McIlwain, D. (2016) "Cognition in skilled action: meshed control and the varieties of skill experience", *Mind and language* 31(1): 37-66.

Christoff, K., Irving, Z., Fox, K., Spreng, N. & Andrews-Hanna, J. (2016) "Mind-wandering as spontaneous thought: a dynamic framework", *Nature reviews neuroscience* 17(11): 718-731.

Chun, M. (2003) "Scene Perception and Memory", in D. Irwin & B. Ross (eds.), *Psychology of Learning and Motivation: Advances in Research and Theory: Cognitive Vision Vol. 42* (pp. 79-108), San Diego: Academic Press.

Clarke, R. (2010) "Skilled activity and the causal theory of action", *Philosophy and phenomenological research* 80(3): 523-550.

Craver, C. (2007) *Explaining the Brain. Mechanisms and the Mosaic Unity of Neuroscience* New York: Oxford University Press.

Cummins, R. (1983) *The nature of psychological explanation* Cambridge, MA: MIT.

Davidson, D. (1970) "Agency", in Davidson 1980, *Essays on actions and events*, Oxford: OUP.

D'Esposito, M. & Postle, B. (2015) "The cognitive neuroscience of working memory", *Annual review of psychology* 66: 115-42.

Dennett, D. (1978) *Brainstorms: philosophical essays on mind and psychology* Montgomery, VT: Bradford books.

Diamond, A. (2013) "Executive functions", *Annual review of psychology* 64: 135-68.

Dixon, M., Fox, K. & Christoff, K. (2014) “A framework for understanding the relationship between externally and internally directed cognition”, *Neuropsychologia* 1-10.

Dombrowe, I., Donk, M. & Olivers, C. (2011) “The cost of switching attentional sets”, *Attention, perception, and psychophysics* 73(8): 2481-2488.

Donoso, M., Collins, A. & Koechlin, E. (2014) “Foundations of human reasoning in the prefrontal cortex”, *Science* 344: 1481-1486.

Dube, B., Basciano, A., Emrich, S. & Al-Aidroos, N. (2016) “Visual working memory simultaneously guides facilitation and inhibition during visual search”, *Attention, perception, and psychophysics* 78: 1232-1244.

Duncan, J. & Humphreys, G. (1989) “Visual search and stimulus similarity”, *Psychological review* 96(3): 433-458.

Duque, J., Olivier, E. & Rushworth, M. (2013) “Top-Down Inhibitory Control Exerted by the Medial Frontal Cortex during Action Selection under Conflict”, *Journal of Cognitive Neuroscience* 25(10): 1634-1648.

Eckstein, M. (2011) “Visual Search: A Retrospective”, *Journal of Vision* 11(4): 1-36.

Embrich, S., Al-Aidroos, N., Pratt, J. & Ferber, S. (2010) “Finding memory in search: the effect of visual working memory load on visual search”, *The quarterly journal of experimental psychology* 63(8): 1457-1466.

Frankfurt, H. (1978) “The problem of action”, in H. Frankfurt (1988), *The importance of what we care about*, CUP.

Fridland, E. (2021) “Intention at the interface”, *Review of philosophy and psychology* 12(3): 481-505.

Fuster, J. (2015) *The prefrontal cortex* New York: Academic Press.

Gazzaniga, M., Ivry, R. & Mangun, G. (2014) *Cognitive Neuroscience. The Biology of the Mind* New York: Norton.

Goldstein, S., Naglieri, J., Princiotta, D. & Otero, T. (2014) “A history of executive functioning as a theoretical and clinical construct”, in S. Goldstein & J. Naglieri (eds.), *Handbook of executive functioning*, Springer: New York.

Hakun, J. & Ravizza, S. (2012) “Cognitive control: preparation of task switching components”, *Brain research* 1451: 53-64.

Han, S. (2015) “Working memory contents revive the neglected, but suppress the inhibited”, *Cognition* 145: 116-121.

Han, S. & Kim, M. (2004) “Visual Search Does Not Remain Efficient When Executive Working Memory Is Working”, *Psychological Science* 15(9): 623-628.

Han, S. & Kim, M. (2009) “Do the contents of working memory capture attention? Yes, but cognitive control matters”, *Journal of experimental psychology* 35: 1292-1302.

Hanna, J., Irving, Z., Fox, K., Spreng, R. & Christoff, K. (2018) “The neuroscience of spontaneous thought: an evolving interdisciplinary field”, in Christoff & Fox (eds.), *The Oxford handbook of spontaneous thought: mindwandering, creativity, and dreaming*, OUP: Oxford.

Hickey, C., Di Lollo, V. & McDonald, J. (2009) "Electrophysiological indices of target and distractor processing in visual search", *Journal of cognitive neuroscience* 21: 760-775.

Hollingworth, A. (2014) "Guidance of Visual Search by Memory and Knowledge", in M. Dodd & J. Flowers (eds.), *The Influence of Attention, Learning, and Motivation on Visual Search*, Nebraska Symposion on Motivation.

Hornsby, J. (1996) *Simple mindedness: in defense of naïve naturalism in the philosophy of mind* Cambridge, MA: Harvard.

Hornsby, J. (2004) "Agency and alienation", in Caro & MacArthur (eds.), *Naturalism in question*, Cambridge, MA: Harvard.

Hyafil, A., Summerfield, C. & Koechlin, E. (2009) "Two mechanisms for task switching in the prefrontal cortex", *J. Neurosci.* 29(19): 5135-5142.

Hyman, J. (2015) *Action, Knowledge, and Will* Oxford: OUP.

Irving, Z. (2016) "Mind-wandering is unguided attention: accounting for the 'purposeful' wanderer", *Philosophical studies* 173:547-571.

Irving, Z. & Glasser, A. (2019) "Mind-wandering: a philosophical guide", *Philosophy compass* 15(1): e12644

Irving, Z. (2021) "Drifting and directed minds: the significance of mind-wandering for mental agency", *Journal of philosophy* 118(110): 614-644.

Itti, L. & Baldi, P. (2009) "Bayesian Surprise Attracts Human Attention", *Vision Research* 49: 1295- 1306.

Itti, L., Koch, C. & Niebur E. (1998) "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20: 1254-1259.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Phillip, A. & Koch, I. (2010) "Control and interference in task switching – a review", *Psychological bulletin* 136(5): 849-874.

Kim. J. (1998) *Mind in a physical world. An essay on the mind-body problem and mental causation* Cambridge, Ma.: MIT Press.

Koechlin, E. (2014) "An evolutionary computational theory of prefrontal executive function in decision-making", *Phil. Trans. R. Soc.* 369: 20130474.

Koechlin, E. & Summerfield, C. (2007) "An information theoretical approach to prefrontal executive function", *Trends in Cognitive Science* 11: 229-235.

Korsgaard, C. (2009) *Self-constitution: agency, identity, and integrity* Oxford: OUP.

Kristjansson, A. & Campana, G. (2010) "Where perception meets memory: a review of repetition priming in visual search tasks", *Attention, Perception & Psychophysics* 72(1): 5-18.

Lavie, N. & Dalton, P. (2014) "The load theory of attention and cognitive control", in A. Nobre & S. Kastner (eds.), *Oxford Handbook of Attention*, Oxford: Oxford University Press.

Leonard, M. & Egeth, P. (2008) "Attentional guidance in singleton search: an examination of top-down, bottom-up, and intertribal factors", *Visual cognition* 16(8): 1078-1091.

Li, L., Wang, M., Zhao, Q. & Fogelson, N. (2012) "Neural Mechanisms underlying the Cost of Task Switching: An ERP Study", *PLOS One 7(7), e42233*: 1-13.

Lu, J., Tian, L., Zhang, J., Wang, J., Ye, C. & Liu, Q. (2017) "Strategic inhibition of distractors with visual working memory contents after involuntary attention capture", *Nature scientific reports* 7: 1-9.

Machamer, P., Darden, L. & Craver, C. (2000) "Thinking about mechanisms", *Philosophy of science* 67(1): 1-25.

Mele, A. (1992) *The springs of action* Oxford: Oxford University Press.

Mertes, C., Wascher, E. & Schneider, D. 2016. "From capture to inhibition: how does irrelevant information influence visual search? Evidence from a spatial cueing paradigm." *Frontiers in human neuroscience* 10: 1-13

Miller, E. & Cohen, J. (2001) "An integrative theory of prefrontal cortex function", *Annual review of neuroscience* 4: 167-202.

Miller, E. & Buschman (2013) "Cortical circuits for the control of attention", *Current opinion in neurobiology* 23:216-222.

Miyake, A., Friedman, N., Emerson, M., Witzki, A., Howerter, A. & Wager, T. (2000) "The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis", *Cognitive Psychology* 41: 49-100.

Mole, C. (2011) *Attention is Cognitive Unison* Oxford: OUP.

Monsell, S. & Driver, J. (2000) "Banishing the control homunculus", in Monsell & Driver (eds.), *Control of cognitive processes XVIII (pp. 2-32)*, Cambridge, MA: MIT Press.

Munakata, Y., Hard, S., Chatham, C., Depue, B., Banich, M. & O'Reilly, R. (2011) "A Unified Frameworke for Inhibitory Control", *Trends in Cognitive Sciences* 15: 453-459.

Murray, S., Irving, Z. & Krasich, K. (forthcoming) "The scientific study of passive thinking: the methodology of mind-wandering research", in DeBrigard & Sinnot-Armstrong (eds.), *Philosophy & Neuroscience*, MIT Press: Cambridge, MA.

Nagel, T. (1986) *The view from nowhere* Oxford: OUP.

Najemnik, J. & Geisler, W. (2009) "A Simple Summation Rule for Optimal Fixation Selection in Visual Search", *Vision Research* 49: 1286-1294.

Nicholson, R., Karayanidis, F., Bumak, E., Poboka, D & Michie, P. (2006) "ERPs Dissociate the Effects of Switching Task Sets and Task Cues", *Brain Research* 1095(1): 107-123.

Norman, D., & Shallice, T. (1986) "Attention to action: Willed and Automatic Control of Behavior", in R. J. Davidson, G. E. Schwartz, & D. Shapiro, (Eds.), *Consciousness and Self-regulation: Advances in Research, Vol. IV (pp. 1-18)*. New York: Plenum Press.

O'Brien, L. (2012) "Deviance and causalism", *Pacific philosophical quarterly* 93(2): 175-96.

Oh, S. & Kim, M. (2004) "The role of spatial working memory in visual search efficiency", *Psychonomic bulletin & review* 11(2): 275-281.

Olivers, C., Meijer, F. & Theeuwes, J. (2006) "Feature-Based Memory-Driven Attentional Capture: Visual Working Memory Content Affects Visual Attention", *Journal of Experimental Psychology: Human*

Perception and Performance 32(5): 1243-1265.

Olivers, C. & Eimer, M. (2011) "On the Difference between Working Memory and Attentional Set", *Neuropsychologia* 49: 1553-1558.

Pereboom, D. (2004) "Is our conception of agent-causation coherent?", *Philosophical topics* 32: 275-86.

Piccinini, G. & Craver, C. (2011) "Integrating psychology and neuroscience: functional analyses as mechanism sketches", *Synthese* 183(3): 283-311.

Prinz, J. (2012) *The Conscious Brain* OUP: Oxford.

Posner, M. & Snyder, C. (1975) "Attention and cognitive control", in Solso, *Information processing and cognition*, Erlbaum.

Ravizza, S. & Carter, C. (2008) "Switching set about task switching: behavioral and neural evidence for distinct forms of cognitive flexibility", *Neuropsychologia* 46(12): 2924-2935.

Rougier, N., Noelle, D., Braver, T., Cohen, J. & O'Reilly, R. (2005) "Prefrontal cortex and flexible cognitive control: rules without symbols", *PNAS* 102(20: 7338-7343.

Sakai, K. (2008) "Task set and prefrontal cortex", *Annu. Rev. Neurosci.* 31: 219-45.

Sawaki, R. & Luck, S. (2010) "On a common neural mechanism for preventing and terminating the allocation of attention", *Journal of neuroscience* 32: 10725-10736.

Sawaki, R. & Luck, S. (2011) "Active suppression of distractors that match the content of visual working memory", *Visual cognition* 19: 956-972.

Sawaki, R. & Luck, S. (2013) "Active suppression after involuntary capture of attention", *Psychological bulletin* 20: 296-301.

Searle, J. (1983) *Intentionality. An essay in the philosophy of mind* Cambridge: CUP.

Shepherd, J. (2014) "The Contours of Control", *Philosophical Studies* 170(3): 395-411.

Shoemaker, S. (1981) "Some varieties of functionalism", *Philosophical topics* 12: 93-119.

Smithies, D. (2011) "Attention is Rational-Access Consciousness", in C. Mole, D. Smithies & W. Wu (eds.), *Attention. Philosophical and Psychological Essays*, OUP: New York.

Soto, D., Humphreys, G., Heinke, D. & Blanco, M. (2005) "Early, Involuntary Top-Down Guidance of Attention from Working Memory", *Journal of Experimental Psychology: Human Perception & Performance* 31(2): 248-261.

Soto, D., Hodsoll, J. Rotshtein, P. & Humphreys, G. (2008) "Automatic Guidance of Attention from Working Memory", *Trends in Cognitive Science* 12(9): 342-348.

Steward, H. (2012) *A Metaphysics of Freedom* Oxford: OUP.

Theeuwes, J. (1991) "Exogenous and endogenous control of attention: the effect of visual onsets and offsets", *Perception & Psychophysics* 49(1): 83-90.

Theeuwes, J. (2010) "Top-Down and Bottom-Up Control of Visual Selection", *Acta Psychologica* 135: 77-99.

Theeuwes, J., Kramer, A. Hahn, S. Irwin, D. & Zelinsky, G. (1999) “Influence of Attentional Capture on Oculomotor Control”, *Journal of Experimental Psychology: Human Perception & Performance* 25: 1595-1608.

Torralba, A., Oliva, A., Castelhano, M. & Henderson, J. (2006) “Contextual Guidance of Attention in Natural Scenes: The Role of Global Features on Object Search”, *Psychological Review* 113(4): 766-786.

Tsotsos, J. & Kruijne, W. (2014) “Cognitive programs: software for attention’s executive”, *Frontiers in psychology* 5: 1-16.

Van Moorselaar, D., Theeuwes, J. & Olivers, C. (2014) “In competition for the attentional template: can multiple items within visual working memory guide attention?”, *Journal of experimental psychology* http://dx.doi.org/10.1037/a0036229 .

Vogel, E., Woodman, G., & Luck, S. (2001) “Storage of features, conjunctions, and objects in visual working memory”, *Journal of experimental psychology* 27: 92-114.

Velleman, D. (1992) “What happens when someone acts?”, in Velleman (2000), *The possibility of practical reason*, Ann Arbor: UMichigan Library.

Vickery, T., King, L., Jiang, Y. (2005) “Setting up the target template in visual search”, *Journal of vision* 5: 81-92.

Vogel, E. & Machizawa, M. (2004) “Neural activity predicts individual differences in visual working memory capacity”, Nature 428: 748-751.

Vogel, E., Mccullough, A. & Machizawa, M. (2005) “Neural measures reveal individual differences in controlling access to working memory”, *Nature* 438: 500-503.

Warden, M. & Miller, E. (2007) “The representation of multiple objects in prefrontal neuronal delay activity”, *Cerebral cortex* 17: 141-150.

Warden, M. & Miller, E. (2010) “Task-dependent changes in short-term memory in the prefrontal cortex”, *Journal of neuroscience* 30(47): 15801-15810.

Walker, R. & McSorley, E. (2008) “The Influence of Distractors on Saccade-Target Selection: Saccade Trajectory Effects”, *Journal of Eye Movement Research* 2(3): 1-13.

Walther, D. & Fei-Fei, Li (2007) “Task-Set Switching with Natural Scenes: Measuring the Cost of Deploying Top-Down Attention”, *Journal of Vision* 7(11): 1-12.

Watzl, S. (2010) *The significance of attention* Columbia University Dissertation

Watzl, S. (2018) *Structuring mind: the nature of attention and how it shapes consciousness* Oxford: Oxford University Press.

Weiskopf, D. (2011) “Models and Mechanisms in Psychological Explanation”, *Synthese* 183(3): 313-338.

Weiskopf, D. (2016) “Integrative Modeling and the Role of Neural Constraints”, *Philosophy of Science* 83(5): 674:685.

Weiskopf, D. (2018) “The Explanatory Autonomy of Cognitive Models”, in M. Kaplan (ed.) *Integrating Psychology and Neuroscience: Prospects and Problems* Oxford: Oxford University Press.

Woodman, G. & Arita, J. (2011) “Direct electrophysiological measurement of attentional templates in visual working memory”, *Psychological science* 22(2): 212-215.

Woodman, G., Vogel, E. & Luck, S. (2001) "Visual search remains efficient when visual working memory is full", *Psychological science* 12: 219-224.

Wu, W. (2014) *Attention* Routledge: London.

Wu, W. (2016) "Experts and Deviants: The Story of Agentive Control", *Philosophy and Phenomenological Research* 93(1): 101-126.

Yantis, S. & Jonides, J. (1990) "Abrupt visual onsets and selective attention: voluntary versus automatic allocation", *Journal of experimental psychology: human perception and performance* 16: 121-134.

Zelinsky, G. (2008) "A Theory of Eye Movements during Target Acquisition", *Psychological Review* 115(4): 787-835.

Zelinsky, G. (2015) "The what, where, and why of priority maps and their interactions with visual working memory", *Annals of the New York academy of science*: 154-164.