Abstract: One of the most influential recent accounts of attention is Wayne Wu’s. According to Wu, attention is selection-for-action. I argue that this proposal faces a dilemma: either it denies clear cases of attention capture, or it acknowledges these cases but classifies many inattentive episodes as attentive.

Word count (including footnotes): 4,477

A dilemma for ‘selection-for-action’

1 Introduction

An important part of our mental life is characterized by what we attend to. It turns out to be difficult, however, to say what attention is. One of the most influential recent accounts of attention is Wayne Wu’s. In a series of papers, Wu has argued for a constitutive connection between attention and action. He maintains that:

(SfA) Necessarily, S attends to X if and only if S selects X for action. (Wu 2014: 96)

The idea that attention is selection for action is attractive.\(^1\) Plausibly, a capacity to attend evolved in order to facilitate (complex, goal-directed) action. Wu’s proposal, however, is much stronger. For an individual to attend to some X, Wu claims, it is necessary that X be selected for an actually occurring action:

(N) If S attends to X, then S selects X for action.

Furthermore, Wu maintains the following sufficient condition for a deployment of attention:

(S) If S selects X for action, then S attends to X.

In what follows I argue that Wu’s proposal either under-generates or over-generates deployments of attention. In section 2 I introduce his selection-for-action account. In section 3 I sketch the dilemma of over- or under-generation that I develop in the subsequent

\(^{1}\) Odmar Neumann and Allan Allport have independently made similar proposals. (cf. Allport 1987; Neumann 1987)
sections: either Wu denies that certain cases of attentional capture constitute deployments of attention; or he classifies many inattentive episodes as attentive. Section 4 describes the relevant cases of captured attention. Other authors have suggested that captured attention can constitute a counterexample to Wu’s account.² So far, the literature lacks a discussion of Wu’s reply. In section 5 I scrutinize this reply and conclude that it is not successful. My main critical point, however, is a different one. In section 6 I argue that even if Wu’s reply to the problematic instances of captured attention were successful, he would commit himself to an account that classifies many inattentive episodes as attentive. In section 7 I conclude that Wu indeed faces the dilemma of over- or under-generation.

2 Selection-for-action

The central notion in Wu’s account of attention is that of selecting a stimulus for action. What is selection-for-action? Wu develops his proposal from a case of ordinary vision-guided bodily action. Suppose that an individual is confronted with two balls that she might kick. Further suppose that the individual is capable of kicking the balls with either her left or her right foot. In order for her to act, the individual must “select” some specific “target” or “input” and “couple” it to some specific “action,” “response,” or “output.”³ The individual faces a “behavior space” of different behavioral options – kicking the first ball with her left foot, or kicking the second ball with her right foot. (Wu 2014: 81) Selection-for-action just is the (result of the) process of coupling a given input with an actual agential response or token action. (Wu 2014: 96)

³ “Input” (and “stimulus”) suggests mental inputs to mental processing. But we normally do not visually attend to mental phenomena. We attend to their distal causes. I will assume that attending to X requires processing a stimulus associated with X, for generating an agentic response. Also, Wu sometimes refers to selection of an action’s target, while at other times, he merely refers to an input to some process. Many actions, such as raising one’s arm or walking about, have no target in any familiar sense. I adopt Wu’s more neutral term input.
Wu makes several qualifications to his initial proposal. First, the relevant kind of “input/output-coupling” does not require selection between several different inputs or outputs. Only one input might be available for coupling with an output. Similarly, Wu does not require that the individual have several different possible outputs (or behavioral options) to select between. The relevant coupling of input to output might occur in a “one-one behavior space so long as the action is something that need not be done.” (Wu 2014: 81)

Second, “the input state … is a personal-level state” (Wu 2014: 82). Personal-level states, for Wu, must be psychological states, states “to which the agent stands in psychological relations, say appropriate perceptual, cognitive, or memory states.” (Wu 2014: 81)

Third, the input’s being “coupled” to an “action,” a “response,” or an “output” cannot merely be the causation of one event by another, but requires (non-deviant) processing of the input. Wu writes that the input must “guide the generation of/inform” the output. (Wu 2011: 100/1; Wu 2014: 79)

3 A dilemma

The account’s plausibility depends on Wu’s conception of an appropriate action, response, or output. If the conception is demanding, we will consider fewer episodes attentive. If the conception is permissive, a wide range of episodes will count as attentive. Wu must tailor his conception of an action, response, or output to precisely the cases in which attention is being deployed.

I will argue that, as things stand, Wu faces a dilemma of over- or under-generation. On one hand, Wu might adhere to a notion of selection-for-action, where action is understood in some intuitive, substantive sense. In this case, both intuition and empirical psychology acknowledge many deployments of attention that Wu will have to deny. I will focus on
certain instances of attention capture. No action needs to ensue from such attentional capture. So Wu has to deny that it constitutes a deployment of attention. But his account then under-generates: it ignores a paradigm attentional phenomenon.

On the other hand, Wu might adopt a notion of response-coupling that accommodates such instances of attention capture. This alternative will force Wu to opt for a very permissive conception of action, response, or output. Wu will have to point us to some type of psychological processing of a stimulus that occurs whenever that stimulus captures attention. But, as I will argue, all the proposed candidates for such response-coupling over-generate attentional episodes. The account, in this case, does not distinguish attentional from non-attentional episodes.

In either case, Wu’s selection-for-action account fails to account for attention’s nature. At a minimum, Wu’s account is underspecified, as it stands.

4 Captured attention without response

Attention can be captured. Suppose that you are looking at the traffic when a bright advertisement flashes into the night sky. The bright light captures your visual attention. You covertly (without moving your eyes) orient attention away from the traffic and to the advertisement. Having attended to the ad for an instant, your attention returns to the traffic.

Certain instances of such captured covert attention constitute prima facie counterexamples to (N). Suppose that you select a stimulus – the bright light or the ad – and attend to it, but you do not select it for any action. In this case attentional selection apparently is not for any response whatsoever. You attend to the ad, then avert your

---

4 An anonymous reviewer rightly points out that these instances of captured attention are not the only cases of attention without a response. An individual may intentionally shift attention to a square on a screen without acting upon the resulting perception.
attention. Is your averting your attention a response to the stimulus? It need not be. You might be annoyed at the ad and therefore shift attention back to the traffic. In many cases some response to the stimulus will, indeed, occur. But you may alternatively just shift attention back next. In this latter case, averting your attention is not a response to the stimulus. No processing of the stimulus occurs, beyond the allocation of attention. Prima facie, then, (N) seems false.

Vision science supports this description of such instances of captured attention. According to vision science, two distinct systems – the endogenous and the exogenous system – orient one underlying psychological processing resource: visual attention. Several observations support this claim. The effects of exogenous and endogenous attention on visual processing are similar. Both increase speed and accuracy of visual processing. (Giordano et al. 2009: 7; Carrasco 2011) Neuroscientific evidence suggests that similar neural mechanisms implement endogenous and exogenous visual attention. (Ling et al. 2014; Beck & Kastner 2009) And indeed, there is evidence that exogenous and endogenous deployments of attention interfere with each other, hence plausibly tap into the very same limited resource. (Busse et al. 2008)

The systems orient visual attention in very different ways, however. The endogenous system orients attention on the basis of individuals’ intentions, goals, and expectations. This system operates slowly – it takes 300 ms to deploy attention – and its activity can be sustained for long periods of time. The exogenous system, on the other hand, responds to physically salient (or evolutionarily important) stimuli. These include abruptly appearing objects, bright colors or lights, as well as salient shapes. The system orients attention rapidly, within 100-120 ms (after which its effect decays). The system’s activity is independent of individuals’ current goals, beliefs, or expectations. Indeed, the system functions to reflexively
orient visual attention as against individuals’ current goals and intentions. Exogenous attention disrupts ongoing goal-directed action. Individuals typically cannot suppress exogenous orientation to a stimulus, even if they know that the stimulus interferes with their ongoing actions, and even if they try to suppress the reflex. (Giordano et al. 2009: 8; Carrasco 2011: 1488) All these facts are deeply engrained in empirical research on visual attention. (Cf. Posner 1980; Carrasco et al. 2004)

Captured attention, of course, in many cases just is orientation by the exogenous system. Captured attention involves the reflex-like orientation of visual processing resources by the exogenous system. Vision science hence supports describing certain cases of captured attention as selection of stimuli that is not for – indeed against – ongoing action.

5 Selection-for-action under-generates

Wu is aware of the difficulty that such cases of captured attention pose. How does Wu respond to this difficulty? Wu faces two options. One the one hand, he might deny that these cases constitute genuine deployments of attention. This option is unattractive. Both intuition and empirical psychology acknowledge the above instances of attention capture as paradigmatic. Intuition offers them as starting points for theorizing about attention. Empirical psychology points to characteristic ways of allocating processing resources in explaining the cases’ being attentive.5 Our account of attention should, by default, cover such paradigm deployments of attention.

5 Other extant accounts of attention can explicate how these are deployments of attention, cf. especially Watzl 2017, Prinz 2014, Mole 2010, or some version of Broadbent 1958. (also fn. 15 below).
Alternatively, Wu might argue that, even in these cases, stimuli are selected for action. Wu chooses this second option. He claims that, as against appearances, even in relevant cases of attention capture, “attention always engages a response.”\(^6\) (Wu 2014: 93)

Note the dialectical situation. Wu proposes that attention, by its very nature, is selection of a stimulus for action. Attention capture, \textit{prima facie}, presents instances of attentional selection that is not for action. Wu should identify psychological episodes that are present in each case of attention capture, episodes for which the attended-to stimulus is selected, and that are plausibly actions. These episodes’ occurrence in each case of attention capture should either be intuitively plausible, or Wu should provide arguments in its favor.

Wu considers three possible kinds of episodes that might have these features. In the case of perceptual attention capture, Wu writes, “you (i) move your body towards the object, orienting towards it; you think about it; you explore it with the other senses.” (Wu 2014: 93) About cases of captured intellectual (non-perceptual) attention Wu writes: “Thoughts and images can also pop into your head as an object pops into your sensory field. … [W]hile the input might be the flashing of a specific visual image, the output would be (ii) the maintaining of that very image.” (Wu 2014: 93) In an earlier paper, Wu similarly writes: “Selection for memory [might be said to be located] within the context of the Many-Many Problem: the items that are stored are located within a mental behavioral space where specific items can be selected for memory.” (Wu 2011: 110) And alternatively: “Let’s return to … the tune that pops into our head. Even in these cases, the stimulus is driving (unintentional) behavior.\(^7\) … The tune that pops into our head is (iii) the generation of a conscious auditory experience that is driven by unintentional selection of stored memory of

\(^6\) Note the move from selection for “action” to engaging a “response.” Not all responses are actions. Wu does not specify the relevant kind of response. I return to this point in section 5 below.

\(^7\) I will address the appeal to ‘unintentional behavior’ below.
music. This latter perhaps identifies a common minimal mental case – the bringing to consciousness of unconsciously stored information.” (Wu 2011: 109; my numbering.)

Is it intuitively plausible that in each case of attention capture, one of these kinds of psychological episode occurs? Does Wu establish their occurrence by way of argument? First, consider group (i). Do we, whenever attention is captured, move our bodies toward the source of the capturing stimulus, orient ourselves toward it, or explore it with other senses? Intuitively, none of these behaviors must follow an episode of attention capture. Vision science emphasizes the possibility of covertly orienting attention toward a stimulus. I hence lay these ways of accommodating attention capture to the side.

Do we always think about a stimulus that captures our attention? It seems possible that we should perceptually attend to a stimulus without also thinking about it. Indeed, the case of capture suggests that the individual need not think about the ad, in order for her visual attention to be captured by it. Vision science’s appeal to visual attention as a perceptual processing resource tends to support this possibility. (Carrasco 2011; Block 2015: 17ff.) Furthermore, animals, such as spiders, exhibit captured attention. But it seems unlikely that they engage in anything like (human-type) thought.8 (Cross & Jackson 2010)

We may not understand thought well enough to adjudicate with certainty whether it constitutes a response to each episode of attentional capture. But the points in the last paragraph cast doubt on this claim’s truth. So the burden is upon Wu to provide independent support for this claim.

---

8 Could Wu propose that, not thought, but some kind of ‘higher’ cognitive processing occurs in each case? Could he thus accommodate attentional capture in animals? Wu would have to develop a sufficiently precise characterization of such cognitive processing. He furthermore would have to make plausible that such processing occurs in each case of attentional capture. But this is precisely the task that we faced at the outset: identifying the kind of psychological episode selection for which constitutes attention. See section 5 below.
Next, consider group (ii). Must individuals always maintain the stimulus they attend to? The case of captured attention suggests the contrary. The individual briefly attends to the stimulus. It might subsequently be erased from her mind. Maybe Wu intends that the attended stimulus must be maintained in working memory? Attended stimuli will often be encoded into both working and long-term memory. They are also more likely to be encoded into fragile visual short-term memory. (Germeys et al. 2010) But we do not know that attention *always* causes encoding of a stimulus into memory. And there is no apparent reason to think that attention *necessarily* causes encoding of a stimulus into memory. The individual might attend to the ad and then instantaneously forget what she saw. So again, there are *prima facie* considerations against Wu's claim.

Maybe Wu is appealing to some other type of maintenance? He would not only have to specify the kind of maintenance at issue. He would also have to provide support for the claim that this maintenance occurs in each case of captured attention. Absent such support, Wu's claim does not convince.

Finally consider group (iii). The relevant attention-generating response to a stimulus here is the transition from previously unconscious information to a phenomenally conscious state, the phenomenally conscious aspect of which consists in what it is like to be in that state. (Block 1995) Does an unconsciously stored stimulus become phenomenally conscious in all instances of attention capture?9 There are widely accepted cases of attention to unconscious stimuli. These stimuli do not, in virtue of attentional selection, become conscious. (Kentridge 2011; Norman et al. 2013) While these results can be disputed, denying them imposes an argumentative burden that has not been discharged. Indeed, Wu apparently

---

9 Not all stimuli that capture attention have been previously stored. Wu should drop the storage requirement in favor of requiring that the stimulus has been unconsciously processed.
accepts these results. (Wu 2011: 111) So we have empirical grounds for rejecting Wu’s appeal to an episode of becoming-conscious in each case of attentional capture.

I have so far not considered the question whether any of these episodes constitute exercises of agency. The opposite seems to be the case. Whether I reflexively orient toward a bright, abruptly appearing light, whether my thinking is grabbed by it, whether I maintain its memory, or whether the unconscious percept of the light becomes conscious – these episodes appear intuitively passive. Thinking, maintaining, and orienting are episodes that we can actively engage in. But when reflexively triggered by captured attention, such episodes appear to be paradigmatically passive episodes. Nor do they have the features that are widely assumed to constitute action. Certainly, as Wu admits, one does not intend to perform them. (Davidson 1970) Nor does one try to orient to, think about, maintain, or consciously experience the stimulus. (O’Shaughnessy 1973) On what other basis, then, are we to recognize the episodes as active? Wu would have to both characterize a plausible alternative notion of action, and support the claim that in each case of captured attention, active orienting, thinking, maintaining, or becoming conscious, occurs.

It hence seems possible for an individual to select a stimulus without further engaging in action, an active response, or indeed any response-type processing of that stimulus to yield an output. Capture provides an example of such attentional selection.

---

10 In captured attention, Wu writes, “the resulting actions are automatic … because none of them are intended.” (Wu 2014: 95)

11 Wu does at one point suggest that any input/output-coupling constitutes action. (Wu 2011: 108) Notice that on this proposal, any psychological processing of a stimulus would constitute agency. The claim appears implausible.

12 Could Wu not maintain that these stimuli are at least available for action? Wu rightly rejects this option. (Wu 2011: 108/9) Many stimuli are processed so as for them to influence – be available for – action. Wu would now have to characterize the kind of selection-for-availability-for-action that covers all and only cases of availability for action of a stimulus that are attentive. An anonymous reviewer suggests that Wu might maintain that the stimuli function to be used in action. When no action ensues, attention does not fulfill its function, on this proposal. But the same points apply. Too many stimuli might be argued to function for use in action. We would now have to characterize a notion of selection-functioning-for-use-in-action that distinguishes between attentive and inattentive processing of stimuli. While both amendments hold more promise for providing a
These observations about capture seem not merely intuitively true. Assumptions deeply engrained in the empirical study of attention support them. Wu does not provide sufficient reason for rejecting these observations. So, Wu’s selection-for-action account of attention plausibly under-generates episodes of attentional selection.

6 Selection-for-action over-generates

Are appeals to the attention-capturing stimulus’ maintenance or memory-encoding, to thinking about it, and to its encoding into consciousness not at least promising? Grant, for the sake of argument, that earlier considerations do not rule out these replies to the difficulty posed by captured attention. It has gone unnoticed that the replies commit Wu to a notion of selection that over-generates attentional episodes. My main critical point is not that Wu cannot accommodate episodes of captured attention. It is rather that Wu faces the dilemma of over- or under-generation.

Most researchers on attention agree that it is a kind of selection. But “attention cannot … be merely selection. After all there are many kinds of selection that do not count as attention.” (Wu 2014: 13) The explanatory power of an account of attention hinges on whether it specifies the right kind of selection. What understanding of the relevant kind of selection does (SfA) offer? Initially, it appeared that selection was tied to action. But we saw

necessary condition on attentional selection, neither appears promising as a sufficient condition on such selection. See fn. 9 and section 5 below.

13 Wu does provide an additional argument. He writes: “I do not see any way to derive selection for action from the concept of attention as the latter is the target of analysis. Rather my strategy will be inference to the best explanation. … There are two main options [as to what is required for processing to be attentive]: selection for action and selection for consciousness. … In subsequent chapters I shall argue for the selection for action account as the best explanation by showing that selection for consciousness does not track attention in the required ways.” (Wu 2014: 91) I can here not discuss Wu’s argument against attention as selection for consciousness.

The argument to the best explanation, however, can only be successful if it considers all plausible candidate-accounts of attention. The argument ignores a range of alternatives: selection for working memory (Prinz 2012) or thought (Smithies 2011), the activity of regulating priority structures (Watzl 2017), cognitive unison (Mole 2010), and access to central processing resources (Broadbent 1958). So Wu’s argument to the best explanation does not succeed.
that the episodes discussed in the previous section no longer merit to be characterized as ‘selection-for-action.’

Is there an alternative notion of a ‘response-coupling’ that Wu might appeal to?\(^{14}\) (Wu 2014: 81) Again, note the dialectical situation. Such ‘response-coupling’ must be both necessary and sufficient for attentional selection. The alternative notion of a ‘response-coupling’ should accommodate attentional capture while still distinguishing attentional from non-attentional selection. Can such a notion be derived from the three kinds of episodes that Wu proposes as possible replies to captured attention? Does selection of any of the three kinds suffice for attention?

Consider again group (i). Is all coupling of a stimulus to thought attentive? Probably not. We plausibly think when we unconsciously, inattentively solve chess problems. Possibly even occurrent conscious thinking can be inattentive, as when I decide what to have for lunch while focusing my attention on my writing. Be that as it may, even if all occurrent thinking must be attentive, it still seems implausible that any coupling of a percept, memory, or emotion to thought constitutes an episode of attentional selection. The visual memory of your mother may be fully unconscious and unattended, but lead you to think about her cooking. In order to distinguish \textit{attentional} coupling of a stimulus to thought, from \textit{non-attentional} coupling to thought, Wu would have to characterize the coupling that yields attention. Not any coupling of a stimulus to thought yields attentional selection.

Next, re-consider group (ii). Is every maintaining a visual image, or every encoding into memory, an episode of attention? Not if the image is maintained in long-term memory. Scores of statistical information and detailed information about scenes are continuously encoded into long-term memory. (Joo \textit{et al.} 2009) Not all this information is attended-to.

\(^{14}\) Cf. fn. 7.
And it is quite possible that encoding into working memory does not require attention either. (Fougnie 2008) Maybe there is another, more ‘online,’ kind of maintenance encoding into which yields attentional selection? Suppose that you are seeing a scene. Aspects of your visual image are attended while others are not. Suppose that you maintain this image in an “online” way. Why should the unattended aspects of the image now become attended? One would expect that there should be a type of “online”-maintenance that is inattentive. Wu should distinguish response-coupling to maintenance that does, from response-coupling that does not yield deployments of attention. Not any coupling of a stimulus to its maintenance yields attentional selection.

Finally, re-consider group (iii). Does all bringing to (phenomenal) consciousness of unconsciously stored information involve attention? Suppose that you consciously remember a visual scene, in all its detail. It seems quite possible that aspects of the initial visual image that were unattended remain so when you remember the scene. If we do not restrict such “bringing to consciousness” to remembered stimuli, then it may turn out that any conscious mental state should be attentive. For, quite possibly, all conscious experience involves prior unconscious processing of information. ¹⁵ But we want to distinguish attentive from inattentive conscious episodes. Not any coupling of a stimulus to consciousness yields attentional selection.

So it seems that none of the three kinds of episodes are sufficient to yield attentional selection. Appeal to these episodes does not yield a notion of response-coupling that could help distinguish attentional from non-attentional selection. ¹⁶

---

¹⁵ Cf. fn. 11. Wu rejects the idea that phenomenal consciousness requires attention. (Wu 2011: 111)
¹⁶ Could Wu postulate a disjunction of these three kinds of episodes as the relevant kind of response? No. The same considerations would apply.
Wu at some point mentions “unintentional behavior” (Wu 2011: 109) as the output of response-coupling. But he does not say what qualifies as unintentional behavior. In what sense can encoding of a stimulus into consciousness, thinking about a stimulus, and encoding of a stimulus into memory all be considered pieces of unintentional behavior? All these processes are, of course, instances of individual-level psychological processing. Should we, then, reformulate (SfA) as:

(SfA*) Necessarily, S attends to X if and only if input X is coupled, through non-deviant psychological processing, to some response.

This proposal even more clearly over-generates attentional episodes. Consider the following examples: An unattended visual percept is encoded into long-term memory. An unattended auditory event causes me to be anxious. My anxiety disposes me to get nervous. A percept of optic flow helps update motor-states that control my balance. A bias, stored in long-term memory, influences my conscious reasoning. A percept of the average size of sets of circles primes me to correctly judge their comparative average size. On (SfA*), individuals would be attending to the visual and auditory objects, the anxiety, to optic flow, the bias, and the statistics of average size, in each of these examples. Any person-level, non-deviant psychological processing would qualify as a deployment of attention. This result fails the project of distinguishing between attentive and non-attentive kinds of response-coupling.

7 Conclusion

I conclude that Wu’s account of attention indeed faces the dilemma of over- or under-generation: either it denies clear cases of attention capture, or it acknowledges these cases but
classifies many inattentive episodes as attentive. Maybe there is a conception of ‘response-coupling’ that neither over-generates, nor under-generates. But it appears that the work of characterizing this conception still remains to be done.

Literature


