Seeing Circles: Inattentive Response-Coupling

Denis Buehler, Institut Jean Nicod, École Normale Supérieure Paris, denis.buehler@ens.fr

Abstract: What is attention? On one influential position, attention constitutively is the selection of some stimulus for coupling with a response. Wayne Wu has proposed a *master argument* for this position that relies on the claim that cognitive science commits to an empirical sufficient condition (ESC), according to which, if a subject S perceptually selects (or response-couples) X to guide performance of some experimental task T, she therein attends to X. In this paper I show that this claim about cognitive science is false. Cognitive science allows for inattentive selection-for-task, or inattentive response-coupling. This means that Wu's account is without independent support.

1. Introduction: What Is Attention?

In a series of papers, Wayne Wu has proposed the following constitutive account of attention:

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

This proposal is revisionary: our intuitive conception of attention allows that we can attend without acting, and act without attending. The burden is hence on Wu to provide independent support for his proposal. Wu's *master argument* in favor of (SfA) relies on a claim about method in cognitive science. According to this claim, cognitive science is committed[1] to the principle:

> (ESC) Subject S perceptually attends to X if S perceptually selects [that is, response-couples] X to guide performance of some experimental task T, i.e. selects X for that task. (Wu 2014: 39)

I will argue that this claim misconstrues the science. Cognitive science firmly entertains the possibility of inattentive response-coupling. Wu's account is therefore without independent support. But since (SfA) is revisionary, we should resist it, until such support has been provided.

In Section 2 I introduce Wu's account. Section 3 describes Wu's argument for the empirical sufficient condition (ESC). In Section 4 I provide a *prima facie* counterexample

[1] What does it mean for cognitive science to be so committed? Wu does not say. I return to his issue in Section 6.

to (ESC), an instance of inattentive task-selection or response-coupling. Let me emphasize, however, that my main point in this paper is *not* that this case constitutes a counterexample to (ESC). My main point is rather that the case illustrates how cognitive science is not committed to (ESC). Section 5 elaborates on this point. Section 6 argues that (ESC) misconstrues commitments in cognitive science. (ESC), and hence Wu's account of attention (SfA), are therefore without independent support. In Section 7 I conclude.

2. Selection-for-Action

Wu claims that attention constitutively consists in the selection of an input for action. What *is* selection for action? Wu develops his proposal from an example 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,' that is, 'couple' it to some specific 'response' or 'output.' The individual faces a 'behavior space' of different perceptual inputs and behavioral outputs—kicking the first ball with her left foot, or kicking the second ball with her right foot (Wu 2014: 81). Attention consists in the coupling of one of the possible inputs to one of the possible outputs.

What is it for an individual to 'couple' inputs to outputs? Wu's remarks on the notion of 'coupling' are sparse, despite the central role it plays in his account. The relevant kind of 'input/output-coupling' does not require active selection or choice (Wu 2016: 8). The coupling need not be intended or otherwise involve intention (Wu 2011a: 107ff.; 2014: 82–83, 92ff.; 2016: 116). It need not be conscious (Wu 2011a: 111; 2014: 150). Coupling does not require selection between several different inputs (Wu 2011b: 53; 2014: 81). Indeed, only one input might be available for coupling with an output (Wu 2011b: 53; 2014: 81). Similarly, the individual need not have several different possible outputs to select between (Wu 2014: 81). The output need not be overt bodily behavior; it might be a psychological state or event (Wu 2014: 81). 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, my emphasis). But the output must constitute some sort of 'response' to the input. A response involves the non-deviant processing of the input-information (Wu 2011a: 100; 2016: 107). Finally, "the input state to coupling is a personal-level state" (Wu 2014: 82).

We thus obtain the following formulation of Wu's proposal (Wu 2016: 108ff.):

> (SfA*) S attends to X just in case one of S's individual-level input states/events representing X is coupled with an individual-level output response.

Is (SfA*) *prima facie* plausible? Several authors have pointed out that common sense offers apparent counterexamples to (SfA*) (Watzl 2011; 2017; Jennings & Nanay 2016; Buehler 2018a; 2018b; Jennings 2020). Let us scrutinize the necessary and sufficient condition embedded in (SfA*) in turn.

How about the necessary condition? Suppose that you walk toward the penalty spot, your attention absorbed by the challenge ahead of you, by the fans' cheering and booing. The referee blows his whistle. Its shrill noise briefly captures your attention. For a split-second, the whistle distracts you from your task. You involuntarily, unintentionally, deploy attention to the sound, before focusing again on the penalty. Captured attention involves attentional 'selection' of a stimulus. But this stimulus need not be coupled to a response. Your *only* response to the stimulus may be your briefly attending to it. So the necessary condition seems *prima facie* false.[2]

Is the sufficient condition plausible? Here, too, we find *prima facie* problems. For consider again your walk toward the penalty spot. Your attention is absorbed by the task of scoring a goal. But in walking, you may also rely on a spatial memory of the goal's location, peripheral vision of the goalie's position, and a proprioceptive sense of your own body's posture. All these input-states inform your walking-behavior. They may well be conscious. They arguably are personal-level mental states (Burge 2010: 369ff.). These

[2] In reply to this kind of case, Wu lowers the requirements on relevant responses. Thus Wu proposes as relevant responses the maintenance of a representational state (Wu 2014: 93), the "bringing to consciousness" of a stimulus (Wu 2011a: 109), the mere "altering [of] our perceptual representations" (Wu 2011a: 105) as well as its encoding into memory (Wu 2011a: 109). Surely, he suggests, even when attention is captured, the stimulus triggers at least a response of these kinds? It is not clear that this must be so. And Wu provides no reason for thinking that it is. More importantly, lowering the requirements on selection will over-generate attentional episodes as per the sufficient condition. See Buehler (2018a) for discussion of this point.

states are thus relevantly coupled to a response. But you do not appear to attend to these states. You attend to the ball, the kick, or the task of scoring a goal. So the sufficient condition seems *prima facie* false, too.

Neither of these considerations is decisive. But they do show that Wu's (SfA*) is a *revisionary* account of attention. Because the account apparently conflicts with common sense, the burden is on Wu to provide independent support for it. Wu's support for (SfA*), in turn, crucially relies on his argument for (SfA*)'s sufficient condition.[3] (Wu 2014: 91, 103) Let us turn to this argument now.

3. The Empirical Sufficient Condition

Wu bases his *master argument* for (SfA*)[4] on the premise that cognitive science commits to the following empirical sufficient condition:

> (ESC) Subject S perceptually attends to X if S perceptually selects [that is, response-couples] X to guide performance of some experimental task T, i.e. selects X for that task. (Wu 2014: 39)

---

[3] Wu does not provide a positive argument for the necessary condition. Instead, he proposes rebuttals of the counterexamples to it. (See Buehler 2018a and fn. 2 above for reasons why these rebuttals do not work.) Wu writes: "I do not see any way to derive selection for action from the concept of attention . . . rather my strategy will be an inference to the best explanation" (Wu 2014: 91). This inference crucially relies on the argument for the sufficient condition.

[4] Wu claims that we can generalize from (ESC) to the sufficient condition in (SfA*). In a first step, he contends: "[g]iven that the behavioral capacities that underwrite performance of experimental tasks are of the sort routinely performed in mundane actions, there is no principled reason to divide experimental tasks from mundane bodily actions such as kicking a ball. This suggests an expanded sufficient condition: If S perceptually selects X for bodily actions, then S perceptually attends to X. . . . the leap from the original empirical sufficient condition . . . is small, significant, and plausible" (Wu 2014: 84). While this step is indeed significant, it is neither small, nor plausible. Specific experimental tasks are expressly designed to investigate attention. Their design requires considerable expertise and relies on background theorizing about plausible characteristics of attention. We cannot generalize from these very specifically constrained circumstances to all bodily action. Wu next proposes to even further generalize, so as to allow that non-perceptual response-coupling in bodily action, as well as response-coupling in mental action, are sufficient for attention. Both generalization-steps are problematic. Neither step is supported by independent argument. For a more extensive discussion of the argument, see Buehler (2018b: §3.3). Wu describes that discussion as revolving around intuitions about cases (Wu 2018: 6–7). This description is incorrect. The discussion rather points out that Wu's revisionary proposal is in need of independent support, which the generalization-argument does not provide. The present paper focuses not on the generalization-steps, but on the main premise—(ESC)—that the argument relies on.

Because cognitive science is thus committed, Wu claims, we have reason to accept the principle. Such support would indeed be powerful. I believe that it would probably outweigh appeals to conflicting observations based in our intuitive conception. In what follows I will argue that the science is not so committed.

Why think that cognitive science is committed to (ESC)? Wu appeals to three classical behavioral paradigms from cognitive psychology in support of this claim: the dichotic listening, visual search, and spatial cueing-paradigms. In each paradigm, subjects are instructed to orient their attention to some perceptual stimulus. In the dichotic listening-paradigm, subjects report auditory stimuli that are played into one ear, while ignoring stimuli played into the other. In visual search, subjects search a display for some specific kind of stimulus and deliver a report or behavioral response to it. In the spatial cueing-paradigm, subjects orient attention to a stimulus presented at some specific location and report on that stimulus. In each paradigm, subjects' correct responses to the stimulus (together with reaction time measures) are taken as evidence that they attentionally selected the stimulus. Wu writes: "For each, there is a well-defined target, reaction to which requires selection of that target to inform the response, whether tracking a conversation in verbal shadowing or examining targets in target detection. . . . There is a general assumption that all experiments on attention hold in using these paradigms:" namely, (ESC) (Wu 2014: 85).

Wu also appeals to research on attention in cognitive neuroscience. To propose neural mechanisms that implement attention, he maintains, neuroscience studies neural activity during behavioral tasks that require deploying attention. On one influential theory, for instance, allocations of spatial attention cause neurons to fire more intensely at attended locations. Researchers support this theory by investigating neural activity during deployments of spatial attention. This investigation relies on behavioral criteria for when spatial attention has been deployed. In effect, neuroscientists rely on behavioral tasks of the kind mentioned in the last paragraph: for instance, they ask subjects to select some stimulus in the context of a spatial cueing-paradigm. By thus relying on behavioral tasks, Wu contends, "neuroscientists must rely on [(ESC)]. . . . It is in light of (ESC) that one can interpret the neural response as "attentional" as opposed to something else" (Wu 2014: 69).

Wu concludes, presumably in an inference to the best explanation, that (ESC) "is built into experimental practice both in psychology and neuroscience. It is the only way to get a handle on attention, and it guides interpretation of the data" (Wu 2014: 72). Therefore, Wu contends, cognitive science is committed to (ESC).

But neither considerations from behavioral psychology nor those from cognitive neuroscience are best explained by such a commitment. According to Wu, in order to investigate neural mechanisms or effects of attention, cognitive neuroscience needs some kind of behavioral handle on when attention has been deployed. But this fact does not support the idea that cognitive neuroscience is committed to (ESC). It would be sufficient for cognitive neuroscience if there were clear, paradigmatic deployments of attention that could serve the investigation of attention's neural underpinnings.[5] Similarly, nothing about the behavioral paradigms favors explanation in terms of a commitment to (ESC). Plausibly, subjects attentionally select a stimulus to complete the tasks embedded in each paradigm. Equally plausibly, behavioral psychologists rely on these tasks to study attention, not because they consider *any* selection-for-task an instance of attentional selection, but rather because the tasks have been carefully designed to involve *attentional* task-selection. Investigating clear instances of attentional selection is entirely consistent with the existence of inattentive task-selection or response-coupling.[6]

This kind of criticism would be more powerful if it could be shown that cognitive science does acknowledge the possibility of inattentive task-selection (or response-coupling). I will next illustrate a large body of research that investigates response-coupling in the absence of attention. I will use this research to argue that there is no commitment in cognitive science that *all* individual-level task-selection or response-coupling constitutes attention.

4. Inattentive Response-Coupling: Seeing Average Circle Size

[5] Wu's argument here presupposes that behavioral psychology implicitly commits to (ESC).
[6] Not only does neither set of considerations support Wu's inference. Reflection on these three behavioral paradigms seems too narrow a basis for an inference concerning *all* studies of attention in cognitive science.

In a recent paper, Joo et al. provide evidence that individuals see, and can judge, the average size of circles in an array, without attentionally selecting the circles (Joo et al. 2009). Have a look at Fig. 1. Which size of the display features the set of circles that are larger, on average?

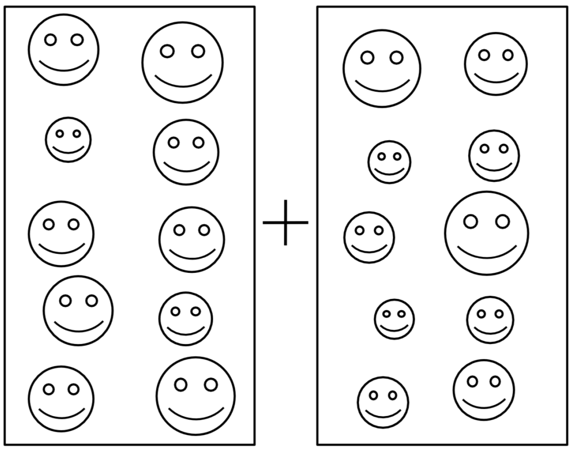

Fig.1 Stimuli from Joo et al. (2009: 2).

The left hand side features circles with an overall greater mean size. You will find—maybe incredibly—that you are extremely good at making such comparative judgments. Most viewers get such judgments right. Joo et al. ask whether seeing average size requires attention to the array of circles. They design an attentional blink paradigm to answer this question. In this paradigm, first introduced by Raymond et al. (1992), a rapid serial visual presentation (or RSVP) stream of visual stimuli is presented to the subjects. Subjects report a first target (T1), as well as (the absence or presence of) a subsequent second target stimulus (T2) in the RSVP stream. (See Fig. 2)

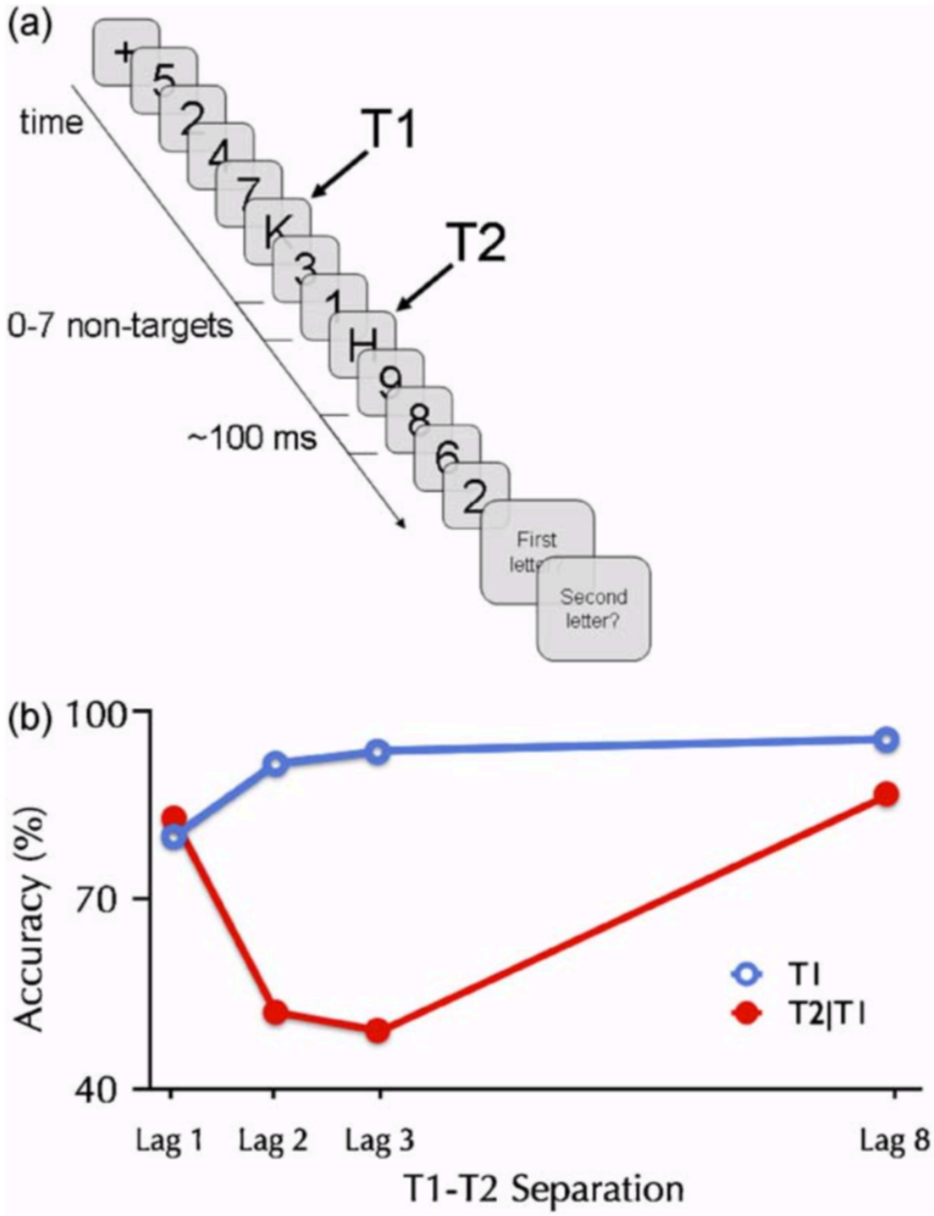

Fig. 2 The basic attentional blink paradigm. (a) A rapid serial visual presentation (RSVP) stream of digits (the non-targets or distractors) is sequentially shown in the middle of the screen, at a rate of 10 items per second. Subjects identify two unspecified letters (targets T1/T2 respectively. The primary measure of interest is the percentage of correct T2 reports from trials in which T1 was correctly identified. (b) Subjects often fail to report T2 when it is presented within 200–500ms after T1. This finding is called "lag 1-sparing." (From Martens & Wyble 2010: 948.)

The finding, replicated in hundreds of experiments, is that if T2 appears between 200–600 ms after T1, subjects typically fail to report T2. The effect basically disappears when subjects ignore T1. On the standard explanation of the effect, subjects cannot report

T2, because they cannot attentionally select it.[7] Researchers refer to the temporal separation between targets as "lag." They found, oddly, that at lag 1 (or if T2 is presented about 100 ms after T1), subjects' report will usually *not* be impaired—a phenomenon called *lag 1-sparing* (Potter et al. 2002).

This is also what Joo et al. find. They use a single digit as their primary target T1. Capital letters constitute distractors. The secondary target T2 is a display of circles similar to that in Fig.1 (see also Fig. 3). Joo et al. manipulate three factors. First, they vary the difficulty of the average size task by manipulating the mean difference of circle-size between the two arrays. Second, they alter the time lag between T1 and T2.[8] Third, they compare a condition in which subjects monitor only T2 with a condition in which they identify both targets. Joo et al. find that task performance is independent of the lag between T1 and T2. Subjects' accuracy is at 69.0% in the condition featuring T1 and T2, and at 68.9% for the condition only featuring T2.

[7] The literature on attentional blink does not always distinguish attention's exogenous, endogenous, spatial, feature-based, and object-based variety. I return to this issue below. Where required I will explicitly point out what kind of attention is at issue.

[8] They test for lag 1, 3, 4, 7, and 10. Each display appears for 59 ms, with a gap of 12 ms between successive displays.

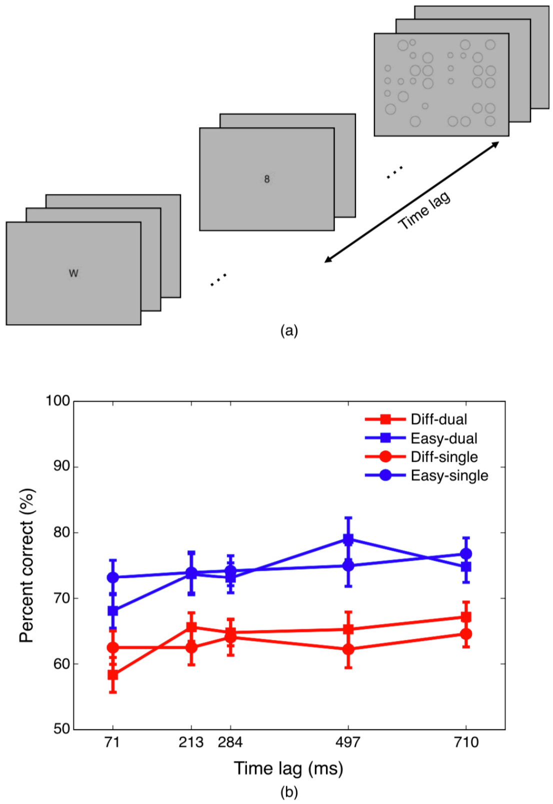

Fig. 3 Experimental setup and results from Joo (2009: 7)

Joo et al. conclude: "evidently the refined stimulus information used for computing mean size remains available even in the absence of focused attention" (Joo et al. 2009: 1).

How do these result bear on (ESC)? They provide *prima facie* evidence that individuals can see, and couple to comparative judgments, the average size of circles in a display, without attentionally selecting the circles. Subjects appear to couple properties of the input visual stimulus—the circles' average size—to the output-response of judging relative average size. Subjects engage in response-coupling of the relevant kind—'selection-for-task.' According to (ESC), such coupling is sufficient for subjects to attend to the input visual stimulus. But in this case, they apparently do so *without attending*. So, (ESC) appears to be false.

How might Wu reply?[9] The most promising reply is that the circles received sufficient attention for being selected and encoded into working memory, after all.[10] To address this reply it will be helpful to delve a little deeper into methodological questions concerning the attentional blink. In the next section I discuss how that literature supports the idea that studies like Joo et al.'s describe instances of inattentive task-selection or response-coupling. But bear in mind throughout that my main aim is *not* to establish that they do. It is rather to show that cognitive science, rather than being committed to (ESC), rejects it. I return to this point in the last section.

5. The Attentional Blink

The effect of the attentional blink, described in the last section, has been known at least since the 1980s (Broadbent & Broadbent 1987).[11] Raymond et al. first introduced the term "attentional blink" in 1992 and proposed the experimental paradigm that has become standard for investigating the effect. (Often, researchers refer to the paradigm, rather than the effect, when they use the term "attentional blink.") Raymond et al. were also the first to suggest, and provide evidence, that this effect is due to constraints on attentional selection (Raymond et al. 1992: 851, 854). They proposed a gating model as explanation of the effect (1992: 858). On this model, visual attention operates like a gate to further processing of the stimuli. When attention selects T1, the gate opens, and an attentional episode is initiated. After around 100 ms, or upon identification of the target, the gate closes and the attentional episode terminates. The resulting inhibition or

[9] Wu might deny that the case exemplifies response-coupling—although it is unclear on what grounds. 'Response-*coupling*' is any non-deviant, individual-level processing of a stimulus-input to some mental or bodily response, as long as it "need not occur" (Wu 2014: 81). Wu does not specify what responses of the appropriate kind are. But all his examples are experimental tasks of precisely the kind at hand. They meet Wu's explicit criteria for (ESC): they are cases "where the subject selects some target to guide their response in carrying out an instructed task" (Wu 2014: 39). Neither the encoding of statistical information from the perception, nor the report of this information "must occur." Indeed, Wu explicitly refers to the kinds of experiments discussed in the main text as instances of response-coupling (Wu 2014: 163–64). They are experiments in which "we have two forms of top-down, goal-directed attention, given task instructions" (Wu 2014: 175).

[10] Wu briefly comments on this research at (Wu 2014: 166–67). He indeed favors the claim that attention is not fully locked to the primary task. Wu appeals to (ESC) to argue that subjects attend to T2. This argument is, of course, not applicable here, since (ESC)'s truth is at issue.

[11] They rely on RSVP paradigms first introduced by Potter and Levy (1969). These paradigms are designed to investigate temporal aspects of information processing in the visual system.

suppression of attentional processing ensures that distractor stimuli will not interfere with target processing, and that hence T1 is accurately reported. The model explains why T2 can typically be identified at lag 1, but not after that.

The following three decades have seen a flurry of follow-up studies, refinements of experimental methods, as well as a range of alternative explanations of the effect. One influential competing account was Chun and Potter's 1995 bottleneck (or two-stage) model of the attentional blink. They proposed, and provided evidence, that stimuli in the attentional blink are rapidly recognized at a first stage of processing—which does not yet require attention. At a second stage, these stimuli must be consolidated and encoded into visual working memory. It is here that attentional selection is required. T2 will not be so encoded when too many attentional resources are still devoted to consolidating T1 for report. Lag 1 sparing is explained as due to T2's benefitting from the same attentional resources allocated to T1.

Yet another account proposes, not that attentional selection itself is suppressed, but rather that control over the attentional filter is disrupted. The first such proposal was made by Di Lollo et al. (2005). On their proposal, attention, in effect, selects the wrong stimuli. At lag 1, T2 is detected, because the attentional filter is still attuned to the correct target. At later lags, attention selects distractors, hence interfering with accurate report. Other versions of this kind of account have it that attention cannot be shifted fast enough from T1 to T2 to enable attentional selection (Jolicoeur et al. 2006).

Note that all these competing accounts explain the attentional blink as due to the absence of attention from T2. They merely differ with respect to what, according to them, causes this absence—suppression of attentional selection, of working memory encoding, or the disruption of attentional control. These early studies were all based on *behavioral* paradigms. Dux and Marois seem to reflect the scientific consensus when they summarize their review of this work on the attentional blink:[12] "We conclude that the AB arises from attentional demands of T1 for selection, working memory encoding, episodic

[12] This review provides an extensive discussion of these, as well as less promising accounts (such as interference accounts, e.g., Shapiro et al. 1994). (See also Zivony & Lamy 2021.) All three proposals have seen a range of variants and extensions. Thus variants of Raymond et al.'s account are: Nieuwenstein (2006), Olivers, Van der Stigchel and Hulleman (2007), Olivers and Meeter (2008), Raffone et al. (2014). For Chun and Potter, see Ward, Duncan, and Shapiro (1996), Jolicoeur (1998), Potter, Straub, and O'Connor (2002). For disrupted-control accounts, see also Taatgen et al. (2009).

registration, and response selection, which prevents this highly-level central resource from being applied to T2 at short T1-T2 lags" (Dux & Marois 2009: 1683).

One may worry, however, that behavioral measures are not fine-grained enough to decide between these accounts, and, more to our point, to ascertain whether or not T2 is attentionally selected. Research soon turned to methods in cognitive neuroscience to further investigate the attentional blink. One particularly powerful method in this context is the recording of event-related potentials (ERPs), or electrical activity from the scalp that is time-locked to specific psychological events and processes (Luck 2014; see Fig. 4). Such studies have the advantage that "they provide a continuous measure of the neural activity that occurs between the onset of the critical event and the response—unlike behavioral measures . . . which reflect the final outcome of the many intervening processes" (Zivony & Lamy 2021: 10). Due to their more precise temporal resolution, ERP-studies are particularly helpful in identifying *what* stages of processing precisely are disrupted during the attentional blink. Maybe not surprisingly, these studies paint a more complex picture than the behavioral studies did.

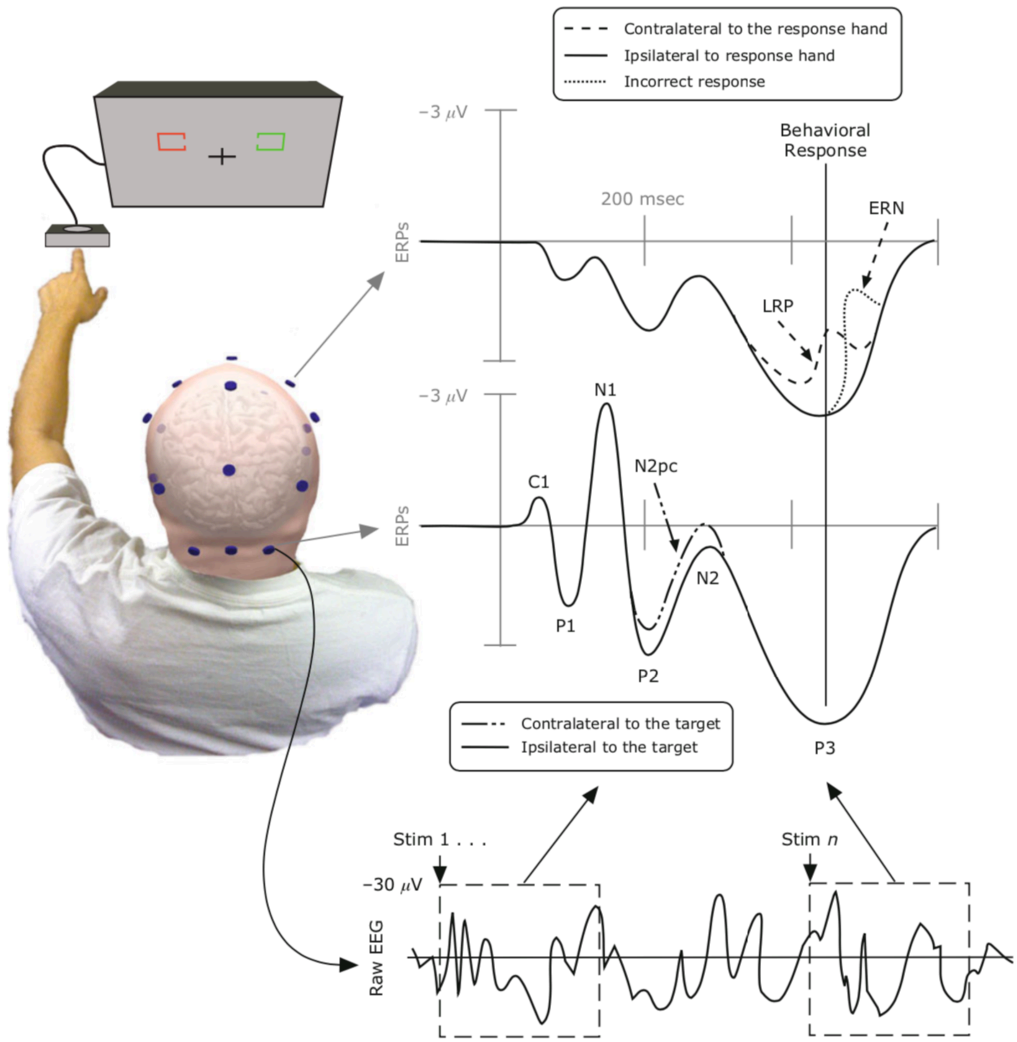

Fig. 4 Set up of an ERP-study, as well as (idealized) components of the ERP-wave, as elicited during a visual search task. The left panel shows a stimulus array to which waveforms on the righthand side are time locked. The bottom right panel illustrates how ERPs are extracted from the raw EEG data. The middle right panel shows the classic sequence of ERP components elicited during a visual task. The top right panel shows the sequence of ERP components observed during the performance of the task requiring a response with a finger on the left hand that could either be correct or incorrect. (From Woodman 2010: 2033.)

Several decades of research have identified correlations between psychological events and components of the ERP-wave that serve as their signatures (Woodman 2010; Luck & Kappenman 2012; Luck 2012 for reviews). For present purposes, only the N2pc and P3 components are relevant. The N2pc component is widely considered a signature of attentional selection of a stimulus (for further processing). Attentional selection of a stimulus fails if the N2pc component is disrupted (Martens & Wyble 2010: 949; Luck

2012: 20ff.; Zivony & Lamy 2016: 1888; Zivony & Lamy 2021: 11). The P3 component indicates the encoding or consolidation of a stimulus in visual working memory. Encoding and maintenance of a stimulus fails if the P3 component is disrupted (Martens & Wyble 2010: 949; Perez & Vogel 2012; Zivony & Lamy 2021: 11).

Dozens of studies have found that, during the attentional blink, both the P3 and N2pc components are disrupted, even entirely absent from the ERP-wave. (Cf. Martens & Wyble 2010; Zivony & Lamy 2021 for reviews.)[13] These results strongly support the picture according to which the attentional blink prevents both attentional selection of T2, and its encoding and consolidation in visual working memory. A particularly striking result from Dell'Acqua et al. (2006) illustrates this work. Their behavioral paradigm was a standard attentional blink. T1 consisted of pairs of digits. Distractor letters followed these digit-pairs. T2 consisted of a colored square with a small gap. Subjects had to report whether the T1 digits were same or different, as well as whether the T2 squares exhibited a gap in the top, bottom, left, or right side. They found that "the N2pc [component of the ERP-wave] was entirely suppressed in the dual-task condition" (Dell'Acqua et al. 2006: 397 and discussion [see Fig. 5]; cf. also Sergent et al. 2005; Jolicoeur et al. 2006; Dell'Acqua et al. 2016; Tang et al. 2020).

[13] Thus write Zivony and Lamy: "The findings converge to show that early perceptual processing (indexed by P1) is unaffected during the blink period, whereas attentional engagement (indexed by N2pc), WM encoding (indexed by P3), and semantic processing (indexed by N400) are disrupted to various degrees" (Zivony & Lamy 2021: 29).

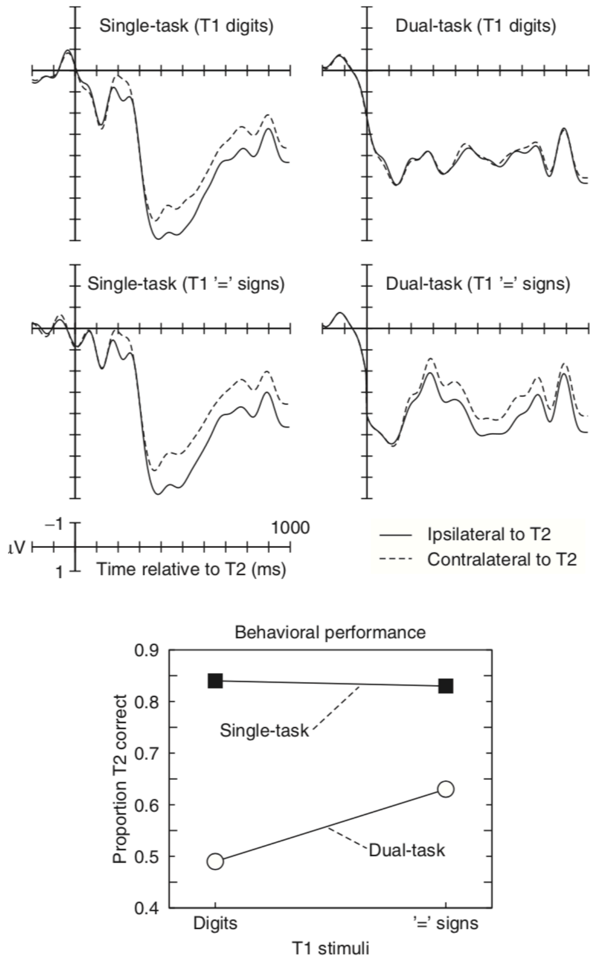

Fig. 5 Behavioral results and ERP wave from Dell'Acqua (2006: 397). Notice that when T1 consisted of digits, we find no N2p-component.

These results support what was suggested by the behavioral studies, that the attentional blink prevents *both* attentional selection and encoding into visual working memory of T2. Indeed, the more fine-grained data from ERP-studies allow a tentative evaluation of the three competing models. Zivony and Lamy suggest, in concluding their review of this research, that the data speak against disruption of control, and in favor of a modified dual-disruption account of attentional selection and working memory encoding:

> These findings suggest that the AB may reflect (i) an early disruption to attentional engagement (with impaired encoding into WM downstream), as suggested by disrupted engagement theories, that occurs only when T1 task is

> difficult enough, and (ii) a later disruption that emerges even with easy T1 tasks and delays WM encoding. (Zivony & Lamy 2021: 36)

Naturally, none of this is beyond dispute. There are different interpretations of different ERP-components (Luck & Kappenberg 2012; Luck 2014). There are results from ERP-studies that are difficult to accommodate for any of the proposed models (Zivony & Lamy 2021). Also, there is the possibility that attention might have effects on psychological processes that are not reflected in the ERP-wave (Luck & Kappenberg 2012: 26). But the consensus appears to be that, during attentional blink, T2 does not receive sufficient attention for it to be attentionally selected.

How does all this bear on the study by Joo et al.? At the end of Section 4 I considered a reply on behalf of Wu according to which information about average circle size received sufficient attention for selection and encoding after all—for instance, because attention was not sufficiently locked at T1 to prevent attentional selection of T2. In light of the consensus on ERP-studies concerning the effect of attentional blink on attentional selection, we have reason to reject this reply. Our most fine-grained measures concur that the attentional blink prevents T2 stimuli from being attentionally selected, identified, and encoded into working memory.

Research on the perception of scene statistics (and scene gist) converges with this result. A wide range of studies suggests that the visual system extracts such statistical information *without* having to rely on attentional selection (Ariely 2001; Chong & Treisman 2003; Cavanagh & Alvarez 2005; Bronfman et al. 2014; Whitney & Leib 2018: §5). Furthermore, there is convergent evidence for this interpretation from other kinds of studies—those involving dual-task paradigms, multiple object tracking paradigms, and so forth (Fei-Fei et al. 2002; 2005; Alvarez & Oliva 2008; Cohen et al. 2011; 2012).

Might the relevant selection not be due to different kinds of attention? Attentional blink and ERP-studies often describe the effect in terms of endogenous visual *spatial* attention. Might the selection be due to exogenous attention, rather than endogenous attention, for instance? One might claim that in Joo et al.'s experiment, subjects fully deploy endogenous attention to T1, but that such allocation does not affect the resources of exogenous attention. Thus while subjects cannot deploy *endo*genous attention to T2, *exo*genous attention may select T2 instead. Subjects' 'selection' of information about circle-size would be attentive after all. However, not only do the stimuli in Joo et al.'s study not appear to be of the kind that would capture exogenous attention—stimuli salient along physical dimensions, such as color and shape singletons, or abrupt onset stimuli. There is also evidence that deploying exogenous attention should cause a decrease in endogenously allocated attention, and hence an effect on T1 processing (Busse et al. 2008). No such effect has been reported.

Another option would be to distinguish between spatial, feature-based, and object-based attention. Absence of the former, one might argue, does not entail absence of either of the latter two. However, it is not enough to merely claim that the results in Joo et al.'s study are due to selection by different kinds of attention. We must *argue* that this is a salient alternative. It seems implausible that object-based attention should be crucially involved in reporting average size. While object-based attention has not been extensively studied in the ERP-paradigm, preliminary data suggests that it relies on spatial attention (Luck 2012: 22). More research has been done on feature-based attention. There is some evidence that this kind of attention can be allocated to T2 during the attentional blink under certain conditions (Zhang & Luck 2009). However, this research also suggests that such feature-based attention does not achieve *selection* of T2 (Tang et al. 2020: 9).

More generally, the literature seems unequivocal in associating attentional selection (for further processing) with the N2pc component—no matter what kind of visual attention is at issue. Different kinds of attention, including spatial visual attention, might be allocated to a stimulus *without this allocation being sufficient for attentional selection* (Luck 2012; Zivony & Lamy 2016; 2018a; 2018b).[14] Any attempt to appeal to different kinds of visual attention as sufficient for selection would hence go against what appears to be the consensus in cognitive science.[15]

---

[14] Suppose that Wu insists that such allocation *is* response-coupling or selection-for-task in the relevant sense. If these authors are right, there is a kind of attentional "selection" that is not followed by any further processing of the selected stimulus. A fortiori, this would be selection without a response in Wu's sense. We would thus have an even more straightforward refutation of Wu's account of attention.

[15] Could there not be yet another kind of attention that does the selecting? Reddy et al. (2007), discussing results from dual task-experiments, suggest allocation of yet some further, unknown alternative attentional resource for face-recognition in dual-task conditions. But a module for face-

This is not to say that this research shows with certainty that no attentional selection occurs for T2 stimuli in Joo's study. I have already commented on disagreements concerning the cognitive significance of different components of the ERP-wave. Could it not be argued that in cases where we find mere disruption, but not full suppression of N2pc, "some" selection occurs? While it is not entirely clear what that might mean, it does seem clear that this is not what the literature assumes. But it may be useful to investigate this possibility in greater depth. It would also be desirable to see the Joo study carried out with an even more cognitively demanding T1 task. Dell'Acqua et al. (2006) found full suppression of N2pc only for a cognitively more demanding T1. T1 in Joo et al.'s study certainly was more demanding that T1 in Dell'Acqua's first, easier condition. But what results would we obtain for the exact same T1 condition for Joo's stimuli? To my knowledge, this experiment has not been carried out so far. Furthermore, one would wish that that very behavioral paradigm be coupled with an ERP-study to ascertain that here, too, the N2pc and P3 components indicating attentional selection and working memory encoding are suppressed, even entirely absent. Again, such a study has, to my knowledge, not yet been conducted. Each of these investigations would strengthen the case made so far.

Even so, a non-biased evaluation of the evidence strongly suggests that, under circumstances as described by Joo, subjects inattentively select average circle-size for task. A range of studies, combining different methods, as well as consensus-interpretation of those methods, thus supports the existence of non-attentional selection-for task.

6. The Empirical Sufficient Condition *Rejected* by Vision Science

Wu's master argument for (SfA*) rests on the claim that cognitive science commits to the following principle:

> (ESC) Subject S perceptually attends to X if S perceptually selects [that is, response-couples] X to guide performance of some experimental task T, i.e. selects X for that task. (Wu 2014: 39)

recognition could equally well explain their results. Furthermore, results about face-recognition do not transfer to other stimuli without argument. And in any case, it is not enough to merely claim that some such attentional selection might be involve without further argument.

Wu does not explain how we can ascertain whether cognitive science is so committed. But I will assume that, if a large number of researchers explicitly allow for the possibility of non-attentional response-coupling, and if there are several, deeply entrenched, widely accepted, and fruitful research paradigms that (implicitly) allow for the possibility of such non-attentional response-coupling, then it is false to say that cognitive science is committed to (ESC).

Joo et al.'s study illustrates both. Joo et al. explicitly allow for the *possibility* of response-coupling of stimuli without attending to those stimuli.[16] Their attentional blink study, as well as the ERP-related work in cognitive neuroscience, illustrates several paradigms that are deeply entrenched in the study of attention, widely accepted, and fruitful. Yet they allow for inattentive response-coupling. Cognitive science thus not only does not support, but rather firmly *rejects* (ESC). Whether or not this research is in the end successful in establishing the existence of response-coupling without attention, it shows that cognitive scientists allow for the in-principle viability of several strategies for studying inattentive response-coupling.

Behavioral paradigms like the attentional blink pursue, roughly, the following strategy for studying inattentive response-coupling:

> Subjects perform task T by response-coupling perceptual information X. By some independent criterion, their attentionally selecting Y prevents them from attentionally selecting X. Performance at T thus exhibits subjects' *inattentive* perceptual response-coupling.

Embedded in this strategy is a commitment to the in-principle possibility of inattentive response-coupling, as well as to the possibility of its empirical study. This commitment does not merely conflict with (ESC): it allows for response-couplings that guide performance of some task, but that do not constitute attentional selection. Acceptance of this paradigm as an in-principle possibility for investigating inattentive vision *also* expresses a commitment to the idea that it is *not* sufficient for attentional selection that individuals couple some stimulus to a response, in the context of performing some task.

[16] See also Fei-Fei (2002; 2005), Alvarez and Oliva (2008), Cohen (2011; 2012).

Research in cognitive neuroscience, such as the study by Dell'Acqua et al., illustrates a neuroscientific criterion for attentional selection. In our case, the criterion consists in the presence of the N2pc-component in the ERP-wave. This criterion is in-principle independent of any behavioral criterion for deployments of attention. Carefully specified behavioral paradigms may, in the long run, merely serve as a heuristic for identifying central, paradigmatic cases of attentional deployments. As pointed out in Section 3, the specification of these paradigms need not be guided by any kind of commitment to a principle such as (ESC). They might rather be guided by the conviction that these paradigms involve attentional selection. Such cases enable the study of attention's neural underpinnings. The neuroscientific project, however, is to find neural signatures and mechanisms of attention. Once such signatures or mechanisms have been found, they may supersede behavioral criteria. They in-principle allow recognizing attentional deployments even if *no* behavioral criterion is met. ERP-readings may not indicate the N2pc-component, even when subjects can report a secondary stimulus, and *vice versa*. The experiment by Dell'Acqua thus illustrates a second strategy for studying inattentive response-coupling:

> Subjects perform task T by response-coupling perceptual information X. Extensive neuroscientific research suggests that behavioral criteria for attentional selection are *merely* heuristic, and should be superseded by neuroscientific criterion NC. Performing task T, subjects do not meet NC; they hence exhibit inattentive perceptual response-coupling.

Embedded in this strategy, too, is a commitment to the possibility of inattentive response-coupling and its empirical study. This strategy, too, not merely conflicts with (ESC), but expresses a commitment to rejecting this principle.

Are the scientists engaging in the research sketched in Section 5 maybe not clear about what they are actually committed to? Are they guilty of some form of irrationality? By no means. It is rather that their commitments concerning attention differ from what is needed to support (SfA*). They typically think of attention as a processing-resource that need not be deployed, even if stimuli are coupled to task-responses. Joo et al., again, explicitly describe their results as showing the "immunity of mean size judgments to depleted attentional resources," due to attention's engagement at primary target (Joo et al. 2009 10).

Paradigms used in studying attention and inattentive vision do not exhibit commitment to (ESC). They rather acknowledge the possibility of response-coupling that is inattentive. Cognitive science thus seems to reject (ESC). A commitment to (ESC) would artificially shut down, from the armchair, fruitful avenues of empirical investigation into attention. Whether or not these paradigms eventually provide evidence for inattentive response-coupling, (ESC) mischaracterizes the methodological assumptions underlying the study of attention.

7. Conclusion

What conclusions does this discussion allow us to draw about attention itself? Wu's *master argument* for (SfA*) centrally relies on his case in favor of (S*). His case rests on (ESC) as its main premise. Support for (ESC), in turn, consists in the claim that cognitive science commits to it. But (ESC) goes *against* commitments in cognitive science. Not only does the argument for (SfA*) fail. (SfA*) conflicts with both common sense and cognitive science. We should resist (SfA*) until independent support has been provided.

Acknowledgments

Thanks to Sam Clarke, Marie Loescher, as well as students in seminars at Oxford and ENS for their feedback. I acknowledge funding from ANR-17-EURE-0017.

References

Alvarez, G. and A. Oliva (2008). The Representation of Simple Ensemble Visual Features Outside the Focus of Attention. *Psychological Science*, *19*(4), 392–401.

Ariely, D. (2001). Seeing Sets: Representation by Statistical Properties. *Psychological Science*, *12*(2), 157–62.

Broadbent, D. and Broadbent, M. (1987). From Detection to Identification: Response to Multiple Targets in Rapid Serial Visual Presentation. *Perception and Psychophysics* 42(2): 105-113.

Bronfman, Z., N. Brezis, H. Jacobson, and M. Usher (2014). We See More Than We Can Report: 'Cost Free' Color Phenomenality Outside Focal Attention. *Psychological Science*, *25*(7), 1394–403.

Buehler, D. (2018a). A Dilemma for "Selection-for-Action". *Thought: A Journal of Philosophy*, *7*(2), 139–49.

Buehler, D. (2018b). Flexible Occurrent Control. *Philosophical Studies*, *176*(8), 2119–37.

Burge, T. (2010). *Origins of Objectivity*. Oxford University Press.

Busse, L., S. Katzner, and S. Treue (2008). Temporal Dynamics of Neuronal Modulation During Exogenous and Endogenous Shifts of Visual Attention in Macaque Area MT. *PNAS, 105*(42)*,* 16380–85.

Cavanagh, P. and Alvarez, G. (2005). Tracking Multiple Targets with Multifocal Attention. *Trends in Cognitive Sciences* 9(7): 349-54.

Chong, S. and A. Treisman (2003). Representation of Statistical Properties. *Vision Research*, *43*(4), 393–404.

Chun, M. and M. Potter (1995). A Two-Stage Model for Multiple Target Detection in Rapid Serial Visual Presentation. *Journal of Experimental Psychology*, *21*, 109–27.

Cohen, M., G. Alvarez, and K. Nakayama (2011). Natural-Scene Perception Requires Attention. *Psychological Science*, *22*(9), 1165–72.

Cohen, M., P. Cavanagh, M. Chun, and K. Nakayama (2012). The Attentional Requirements of Consciousness. *Trends in Cognitive Sciences*, *16*(8), 411–17.

Dell'Aqua, R., P. Sessa, P. Jolicoeur, and N. Robitaille (2006). Spatial Attention Freezes During the Attentional Blink. *Psychophysiology*, *43*, 394–400.

Dell'Acqua, R., M. Doro, P. Dux, T. Losier, and P. Jolicoeur (2016). Enhanced Frontal Activation Underlies Sparing from the Attentional Blink: Evidence from Human Electrophysiology. *Psychophysiology*, *53*, 623–33.

Di Lollo, V., J. Kawahara, G. Shahab, and J. Enns (2005). The Attentional Blink: Resource Depletion or Temporary Loss of Control. *Psychological Research*, *69*, 191–200.

Dux, P. and R. Marois (2009). The Attentional Blink: A Review of Data and Theory. *Attention, Perception, Psychophysics, 71*(8)*,* 1683–700.

Fei-Fei, L., R. VanRullen, C. Koch, and P. Perona (2002). Rapid Natural Scene Categorization in the Near Absence of Attention. *PNAS*, *99*(14), 9596–601.

Fei-Fei, L., R. VanRullen, C. Koch, and P. Perona (2005). Why Does Natural Scene Categorization Require Little Attention? Exploring Attentional Requirements for Natural and Synthetic Stimuli. *Visual Cognition*, *12*(6), 893–924.

Jennings, C. (2020). *The Attending Mind*. Cambridge University Press.

Jennings, C. and B. Nanay (2016). Action Without Attention. *Analysis*, *76*(1), 29–36.

Jolicoeur, P. and R. Dell'Acqua (1998). The Demonstration of Short-Term Consolidation. *Cognitive Psychology*, *36*, 138–202.

Jolicoeur, P., P. Sessa, R. Dell'Acqua, and N. Robitaille (2006). On the Control of Visual Spatial Attention:

Evidence from Human Electrophysiology. *Psychological Research*, *70*, 414–24.

Joo, S., K. Shin, and S. Chong (2009). On the Nature of the Stimulus Information Necessary for Estimating Mean Size of Visual Arrays. *Journal of Vision*, *9*(9), 1–12.

Luck, S. (2012). Electrophysiological Correlates of the Focusing of Attention Within Complex Visual Scenes: N2pc and Related ERP Components. In S. Luck and E. Kappenman (Eds.), *Oxford Handbook of Event-Related Potential Components* (329–60). Oxford University Press.

Luck, S. (2014). *An Introduction to the Event-Related Potential Technique*. MIT Press.

Luck, S. and S. Hillyard (1994). Electrophysioloigical Correlates of Feature Analysis During Visual Search. *Psychophysiology*, *31*(3), 291–308.

Luck, S. and E. Kappenman (Eds.) (2012). *Oxford Handbook of Event-Related Potential Components*. Oxford University Press.

Martens, S. and B. Wyble (2010). The Attentional Blink: Past, Present and Future of a Blind Spot in Perceptual Awareness. *Neuroscience and Biobehavioral Reviews*, *34*, 947–57.

Nieuwenstein, M. (2006). Top-Down Controlled, Delayed Selection in the Attentional Blink. *Journal of Experimental Psychology*, *32*, 973–85.

Olivers, C. and M. Meeter (2008). A Boost and Bounce Theory of Temporal Attention. *Psychological Review*, *115*, 836–63.

Olivers, C., S. Van der Stigchel, and J. Hulleman (2007). Spreading the Sparing: Against a Limited-Capacity Account of the Attentional Blink. *Psychological Research*, *71*(2), 126–39.

Perez, V. and S. Vogel (2012). What ERPs Can Tell Us about Working Memory. In S. Luck and E. Kappenman (Eds.), *Oxford Handbook of Event-Related Potential Components* (329–60). Oxford University Press.

Potter, M. and Levy, E. (1969). Recognition Memory for a Rapid Sequence of Picture. *Journal of Experimental Psychology* 81(1): 10-15.

Potter, M., A. Straub, and D. O'Connor (2002). The Time Course of Competition for Attention: Attention is Initially Labile. *Journal of Experimental Psychology*, *28*, 1149–62.

Raffone, A., N. Srinivasan, and C. van Leeuwen (2014). The Interplay of Attention and Consciousness in Visual Search, Attentional Blink, and Working Memory Consolidation. *Philosophical Transactions of the Royal Society*, *369*, 20130215.

Raymond, J., K. Shapiro, and K. Arnell (1992). Temporary Suppression of Visual Processing in an RSVP Task: An Attentional Blink? *Journal of Experimental Psychology*, *18*, 849–60.

Reddy, L., Reddy, L., and Koch, C. (2007). Face Identification in the Near-Absence of Focal Attention. *Vision Research* 46: 2336-2343.

Sergent, C., S. Baillet, and S. Dehaene (2005). Timing of the Brain Events Underlying Access to Consciousness During the Attentional Blink. *Nature Neuroscience*, *8*, 1391–400.

Shapiro, K., J. Raymond, and K. Arnell (1994). Attention to Visual Pattern Information Produces the Attentional Blink in RSVP. *Journal of Experimental Psychology*, *20*(2), 357–71.

Taatgen, N., I. Juvina, M. Schipper, J. Borst, and S. Martens (2009). Too Much Control Can Hurt: A

Threaded Cognition Model of the Attentional Blink. *Cognitive Psychology*, *59*, 1–29.

Tang, M., L. Ford, E. Arabzadeh, J. Enns, T. Visser, and J. Mattingley (2020). Neural Dynamics of the Attentional Blink Revealed by Encoding Orientation Selectivity During Rapid Visual Presentation. *Nature Communications*, *11*(434). https://doi.org/10.1038/s41467-019-14107-z

Ward, R., J. Duncan, and K. Shapiro (1996). The Slow Time-Course of Visual Attention. *Cognitive Psychology*, *30*(1), 79–109.

Watzl, S. (2011). The Nature of Attention. *Philosophy Compass*, *6*(11), 842–53.

Watzl, S. (2017). *Structuring Mind: The Nature of Attention and How It Shapes Consciousness*. Oxford University Press.

Whitney, D. and A. Leib (2018). Ensemble Perception. *Annual Review of Psychology*, *69*: 105-129.

Woodman, G. (2010). A Brief Introduction to the Use of Event-Related Potentials in Studies of Perception and Attention. *Attention, Perception & Psychophysics*, *72*(8), 2031–46.

Wu, W. (2011a). Attention as Selection for Action. In Mole, C., Smithies, D., and Wu, W. (Eds.), *Attention: Philosophical and Psychological Essays* (97–116). Oxford University Press.

Wu, W. (2011b). Confronting Many-Many Problems: Attention and Agentive Control. *Nous*, *45*(1), 50–76.

Wu, W. (2014). *Attention.* Oxford University Press.

Wu, W. (2016). Experts and Deviants: The Story of Agentive Control. *Philosophy and Phenomenological Research*, *92*(2), 101–26.

Wu, W. (2019). Action Always Involves Attention. *Analysis*, *79*(4), 693–703.

Zivony, A. and D. Lamy (2016). Attentional Capture and Engagement During the Attentional Blink. *Journal of Experimental Psychology*, *42*(11), 1886–902.

Zivony, A. and D. Lamy (2018a). Dissociating between the N2pc and Attentional Shifting: An Attentional Blink Study. *Neuropsychologia*, *121*, 153–63.

Zivony, A. and D. Lamy (2018b). Perceptual Processing Is Not Spared During the Attentional Blink. *Journal of Cognition*, *1*(1), 18, 1–14.

Zivony, A. and D. Lamy (2021). What Processes Are Disrupted During the Attentional Blink? An Integrative Review of Event-Related Potential Research. *Psychonomic Bulletin & Review* 29(2): 394-414.

Zhang, W. and Luck, S. (2009). Feature-based Attention Modulates Feedforward Visual Processing. *Nature Neuroscience* 12(1): 24-25.