Cambridge University Press
978-1-108-79621-7 — The Breadth of Visual Attention
Stephanie C. Goodhew
Excerpt
More Information

1 Introduction to Attentional Breadth

For most people, vision is the primary sensory modality. Vision allows us to navigate through the world and interact with it. It is our means of driving safely through traffic, avoiding obstacles, perceiving food we want to eat, reading, and recognising the face of a loved one. But at any given moment, there is far more information available to process in visual scenes than our brain is capable of processing to the level of awareness. This means that visual attention has a fundamental triaging role to play in shaping our perception of the world, by selecting certain relevant or salient information for privileged processing, while filtering out other information.

There are many different ways in which humans can regulate their visual attention. The metaphor of a ‘spotlight’ of attention has been used (Posner, Snyder, & Davidson, 1980), to convey the notion of a relatively small island of the visual field that is the focus of attention at any moment in time. This implies an enhanced region of processing, to the exclusion of locations or objects. Of course, this spotlight metaphor is imperfect as a model of attention, because this privileged region does not have a sharp edge but tends to gradually decline from the central focus (Downing, 1988; Eriksen & St. James, 1986; White, Ratcliff, & Starns, 2011). This region sometimes appears to have a non-monotonic roll-off function, such that the intensity of the focus does not just gradually decrease with increasing distance from central focus, but can instead reverse in direction of change (e.g., increase/decrease) of intensity with increasing distance (Caparos & Linnell, 2010; Cutzu & Tsotsos, 2003; Mounts, 2000a, 2000b; Müller, Mollenhauer, Rösler, & Kleinschmidt, 2005), rather than having a ‘hard’ edge as a spotlight might suggest. However, it is a useful metaphor in that it conveys the main idea of a locus of spatial attention, that is distinct from the notion of attention being applied uniformly across the visual field.

Humans can regulate their visual attention in many ways. For example, the central focus of attention (i.e., spotlight) can be shifted (translated) across space (Petersen & Posner, 2012; Posner, 1980). This spotlight does not have to be singular but can instead be split into multiple non-contiguous locations (Müller, Malinowski, Gruber, & Hillyard, 2003). It can also take on shapes other than a circle or ellipse, such as an annulus (doughnut) shape (Bleckley, Durso, Crutchfield, Engle, & Khanna, 2003; Jefferies & Di Lollo, 2015). This Element is focussed on attentional breadth, it considers the possibility that the size of the attentional spotlight can be contracted and expanded. This process has been likened to a zoom lens of a camera, in that there is a tradeoff between the area of focus and the resolution of this focus (Eriksen & St. James, 1986). When a zoom lens is narrowly focussed, only a small

Cambridge University Press
978-1-108-79621-7 — The Breadth of Visual Attention
Stephanie C. Goodhew
Excerpt
More Information

region of a scene is visible in sharp detail, when it is expanded a wider field of view becomes visible but at the expense of perceptual resolution. (This model will be discussed further in Section 3). This Element will discuss the nuances of conceptualising, measuring and manipulating attentional breadth, and it will review the theoretical development that has occurred in our understanding of the consequences of maintaining a given attentional breadth (e.g., broad versus narrow) for performance on visual tasks. Finally, it will consider the utility of the distinction between exogenous and endogenous attentional orienting in relation to attentional breadth.

Given that this Element is about attentional breadth, the first question we should address is its definition. What does it mean to have a narrow versus broad focus of attention? These are non-trivial questions. Some authors offer a concrete definition of attentional breadth, such as ‘the visual area in which information can be acquired within one eye fixation’ (Ball, Beard, Roenker, Miller, & Griggs, 1988). Others have used definitions that are more tied to the specific stimuli used, such as the processing of global versus local elements (Dale & Arnell, 2013). Some eschew a definition altogether. Attentional breadth-related concepts go by many different names in the literature, including attentional spotlight, attentional scale, attentional scope, attentional spread, attentional window, attended region size, useful field of view, and global versus local bias (Balz & Hock, 1997; Bulakowski, Bressler, & Whitney, 2007; Chong & Treisman, 2005; Dale & Arnell, 2015; Edwards, Fausto, Tetlow, Corona, & Valdes, 2018; Fang et al., 2017; Fang, Sanchez-Lopez, & Koster, 2018; Goodhew & Edwards, 2016; Goodhew, Shen, & Edwards, 2016; Heitz & Engle, 2007; Huttermann, Memmert, & Simons, 2014; Kosslyn, Brown, & Dror, 1999; Lawrence, Edwards, & Goodhew, 2020). Here, at the outset, all of these will be treated as belonging to the broad umbrella term of attentional breadth. As will become apparent in the discussion in Section 2, there are likely at least several important sub-processes or subtypes of attentional breadth. However, I will allow this to emerge organically from the review, rather than pre-empt them with strict definitions. For now, the working definition will be that attentional breadth refers to the spatial extent of the area over which spatial attention is applied, with the assumption that the area is contiguous and approximately elliptical. Throughout this Element, I will refer to manipulations that induce small (or narrow) versus large (or broad) attentional breadths. To clarify, by using these terms, I am not invoking absolute categories, but instead referring to relative sizes along a continuum of attentional breadth. That is, the effect of a given attentional breadth can only be compared with that of other attentional breadths, there is no reason to favour the notion of an absolute value of a large versus a small one.

Cambridge University Press
978-1-108-79621-7 — The Breadth of Visual Attention
Stephanie C. Goodhew
Excerpt
More Information

Of course, another fundamental and important question is, why attentional breadth? Why is this Element devoted to this issue? When considering the attention literature, there is a larger amount of literature on shifts of attention, including the factors that regulate them, and their performance consequences for visual tasks, compared with that for attentional breadth. Attentional breadth is critically important. This is because the size of the attentional breadth can alter even what would be typically considered fundamental visual processes, such as our spatial or temporal acuity (Lawrence, Edwards, & Goodhew, 2020; Mounts & Edwards, 2017). That is, adopting a narrow attentional breadth can improve the level of fine spatial detail that we can resolve, such as detecting the presence of a small spatial gap in a ring. Adjusting the size of the attentional breadth can improve visual processes including those invoked in the opening of this section, such as recognising a person's face (Gao, Flevaris, Robertson, & Bentin, 2011). In addition, broadened attentional breadth has been found to be related to important functional outcomes such as drivers' crash risk (Ball, Owsley, Sloane, Roenker, & Bruni, 1993; Wood, Chaparro, Lacherez, & Hickson, 2012) and the breadth-of-attention is said to reflect a person's emotional and motivational state (Fredrickson & Branigan, 2005; Gable & Harmon-Jones, 2010a) and even regulate mood (Gu, Yang, Li, Zhou, & Gao, 2017). Finally, different measures of attentional breadth have been found to vary as a function of different individual characteristics, such as age, personality, and working memory capacity (Kreitz, Furley, Memmert, & Simons, 2015; Lawrence, Edwards, & Goodhew, 2018; Wilson, Lowe, Ruppel, Pratt, & Ferber, 2016). It is, therefore, clearly important for us to understand this process, from both a theoretical and a practical perspective.

This Element is timely because the field is at a critical juncture. That is, the field has amassed sufficient evidence to highlight how greater clarity and consensus is required in both the operationalisation and conceptualisation of attentional breadth if the field is to advance. The goal of this Element is to provide a discussion and a framework to guide this process forward.

2 Measuring Attentional Breadth

In this section, I will discuss some of the most commonly used methods designed to measure attentional breadth. In doing so, I will highlight where there are alternative attentional processes that could underlie performance on these tasks, to determine whether they are truly gauging attentional breadth or potentially some other attentional or cognitive process. I will also consider the potential underlying structure of attentional breadth. That is, attentional breadth may not be a monolithic construct. Take, for analogy, working memory. Working memory is an important construct, but it has multiple meaningful subcomponents, such as a central executive, visuospatial sketchpad, and phonological loop (Baddeley, 2012; Baddeley & Hitch, 1974). In the domain of attention, Corbetta and Shulman (2002) differentiate between top-down attentional processes that serve functions including goal execution and action selection, and more bottom-up mechanisms driven by stimulus salience. It is possible, probable even, that attentional breadth may have a multifaceted underlying structure of subcomponents. If true, then this should be revealed via convergent and divergent patterns of associations, in individuals' performance on these tasks. This would be determined by having the same group of individuals perform each of these tasks, and examining the correlation between them (e.g., if individual X is gauged to have a broad attentional breadth on task A and individual Y a narrow breadth on task A, then convergent evidence for tasks A and B would be if individual X was also gauged to have a broad attentional breadth on task B and individual Y a narrow one). That is, theoretically, if attentional breadth is a construct that actually does underlie performance on all of these different tasks, then performance on them should be correlated. In contrast, if there are meaningful distinctions between the different aspects of attentional breadth, then performance on particular types of tasks may diverge from others. For example, while all measures of attentional breadth should have some relationship with one another, if two tasks gauge the same subcomponent of attentional breadth, then their correlation should be higher than either of their correlations with the measure of a different subcomponent of attentional breadth. Other additional types of evidence will be discussed later in this section. Considering alternative explanations for performance on tasks thought to measure attentional breadth is important. If a task does not measure what it claims to, then this could lead to a wrongful conclusion regarding the true structure of attentional breadth. Following the introduction of each of these methods, I will introduce some key criteria to consider in the search for validating evidence for measures of attentional breadth. Then, I will critically discuss the evidence regarding whether or not they do all indeed reflect attentional breadth, and the same aspect of it. However, the available evidence is limited and incomplete, therefore, I will make recommendations regarding the steps that are required in future research to ensure that we have a solid foundation from which to study and understand attentional breadth and its potentially multifaceted nature.

2.1 Navon

Navon stimuli are compound stimuli, in which information can be presented independently to participants at both a global and a local level (see Figure 1

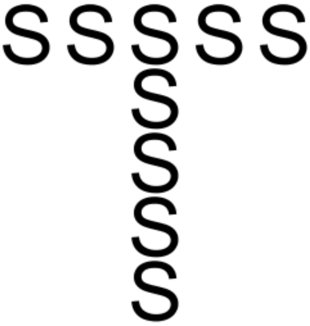

Figure 1 An example illustration of a Navon stimulus. Here, the global letter is ‘T’, whereas the local elements are the letter ‘S’ (note that all figures are intended for the purposes of illustration and are not necessarily to scale).

for an example). The original study (Navon, 1977) used letter stimuli, whereby a larger letter (global level) was made up of multiple occurrences of the same small letter (local level). These stimuli were introduced with the aim of studying the stages of visual scene analysis (Navon, 1977, 1981). It was found that information at the global level was processed and affected performance even when participants were instructed to attend to the local level, whereas the reverse did not occur. The local level did not interfere with the processing of the global level content (Navon, 1977). This led to the conclusion that global processing was completed before local processing, called the ‘global precedence’ (Navon, 1981). Metaphorically, this means that participants are inclined to see the *forest* before the *trees.* The relative advantage for processing the global relative to the local elements of such stimuli is one that has persisted through decades of research (Badcock, Whitworth, Badcock, & Lovegrove, 1990; Baumann & Kuhl, 2005; Hoar & Linnell, 2013; Navon, 1977, 1981).

It should be noted that the concept of global precedence has been challenged (Kinchla & Wolfe, 1979) and that whether or not a processing advantage for the global elements of such stimuli is observed does depend on a number of factors, including the respective sizes of the global and local elements, and how densely the local elements are arranged (Goodhew & Plummer, 2019; Kimchi & Palmer, 1982; Pomerantz, 1983; Yovel, Levy, & Yovel, 2001). Indeed, under certain conditions, a bias in favour of *local* processing can be obtained (Kinchla & Wolfe, 1979). In addition, it has been found that such stimulus parameters have a greater effect on the visual search for global targets than for local targets (Enns & Kingstone, 1995). However, more recent work has supported Navon’s fundamental idea of visual processing progressing from a broad brushstroke initial sweep, followed by the more detailed processing of individual elements (Bar et al., 2006; Greene & Oliva, 2009).

Since their introduction, Navon stimuli have been developed as measures of attentional breadth. There are two main types of tasks where Navon stimuli are used (hereafter, 'Navon tasks' or 'versions of Navon task'). Here, I will refer to them as the *directed* and *undirected* versions of the Navon task. The latter is sometimes called the *divided attention* Navon task, however, I think it is preferable to make as few assumptions as possible about the particular attentional processes that are occurring when naming tasks without evidence to support this. Instead, the directed versus undirected distinction refers to the task instructions and does not make assumptions about whether attention is divided, or instead rapidly switches between the different levels. In the directed Navon task, participants' attention is directed to a particular level of the stimulus using direct task instructions. For example, participants are instructed to identify the letter presented at the global level for a block of trials. At the (task-irrelevant) local level, stimuli that are congruent versus incongruent with the target stimuli at the global level are presented, and participants' response efficiency[1] to identify the global target is compared for the congruent versus incongruent trials. In another block of trials, participants would be instructed to identify the letter presented at the local level, and their relative performance in doing so when congruent versus incongruent information is presented at the global level is gauged. If the interference from the incongruent (relative to the congruent) trials is greater when the target level is at the local level, then this is said to show a global advantage and, therefore, a broad attentional breadth, whereas if interference from the incongruent trials is greater when the target level is the global level, then this is said to show a local advantage and, therefore, a narrow attentional breadth (Caparos, Linnell, Bremner, de Fockert, & Davidoff, 2013; Dale & Arnell, 2013; Navon, 1977).

In contrast, in the undirected version of the Navon task, participants are not instructed to identify the letters at a prescribed level. Instead, they are instructed to identify one of a prescribed set of targets (e.g., the letter 'T' or 'H'). One and only one of these targets appear in each Navon stimulus, and the target can occur at either the global or local level. Participants are simply asked to identify which target stimulus is present, as quickly and as accurately as possible, irrespective

---

[1] For such tasks, by design accuracy is typically at or near maximum (i.e., ceiling). Accuracy is measured to ensure the participants' ability and willingness to comply with task instructions, and to check for speed-accuracy tradeoffs. In such designs, response speed is the primary dependent variable. I use the term 'response efficiency' or 'response speed' to refer to faster responses, assuming equally accurate or more accurate responses in this condition relative to the slower condition.

of which level it appears. Here, the relative response efficiency to the targets at the global versus local level provides an index of attentional breadth. If responses are facilitated for the global relative to the local level, then this is used to infer a relatively broad attentional breadth, whereas if responses are facilitated for the local versus the global level, then this leads to an inference of a relatively narrow attentional breadth (Gable & Harmon-Jones, 2008; Goodhew & Plummer, 2019; McKone et al., 2010). The most common instantiation of the undirected Navon task is to have the target appearing equally often at each level over the block of trials, and trial types are randomly intermixed.

The Navon task has been adapted to measure *attentional resizing efficiency*, which is an important attentional function. This is because different perceptual and cognitive processes benefit from different attentional breadths. For example, if one is trying to resolve final spatial detail in order to read text, then this perceptual process would benefit from a relatively narrow attentional breadth, and correspondingly be compromised by a broad attentional breadth (Balz & Hock, 1997; Lawrence, Edwards, & Goodhew, 2020). In contrast, if one is scanning a crowd to locate a friend, then this would likely benefit from a broader attentional breadth and would be relatively impaired with a narrower attentional breadth (Gao et al., 2011; Macrae & Lewis, 2002). While in the laboratory we typically study the impact of attentional breadth on one process at a time, in real-world vision, humans are often juggling and rapidly switching between multiple tasks with different demands. For example, when driving a car, reading the speedometer requires a narrow focus of spatial attention, in order to resolve the fine spatial information and avoid interference from surrounding instruments, whereas monitoring the road for any change or movement (e.g., a child approaching the road, or the trajectories of other cars) requires a broad focus. Similarly, when meeting up with a friend, one may have to switch between a narrow focus for reading a text message and a broad focus for identifying the friend's face in a crowd, or identifying what direction the bulk of the crowd is moving in. The laboratory research tells us that optimising attention to facilitate these processes requires adopting different attentional breadths. These functional requirements of real-world vision demand that humans can efficiently change between one attentional breadth and another. In other words, they can rapidly and dynamically *resize* their attentional breadth.

Some early literature examined attentional breadth resizing via what was dubbed 'level readiness'. Ward (1982) found that when participants were directed to identify the information at different levels (global/local) of Navon stimuli, they were quicker to do this if the previous trial had required the response at the same level compared with a response about the previous level. Other studies have observed similar results (Robertson, Egly, Lamb, & Kerth,

Cambridge University Press
978-1-108-79621-7 — The Breadth of Visual Attention
Stephanie C. Goodhew
Excerpt
More Information

1993; Wilkinson, Halligan, Marshall, Büchel, & Dolan, 2001). This effect has been observed when the stimulus changes location, colour, polarity, or contrast between trials, and it can persist for up to three *seconds* between one trial and the next (Robertson, 1996). Some work has suggested that even providing participants with a cue about the nature of the upcoming stimulus does not allow them to recover from the deleterious effect of an opposite-level preceding trial (Hubner, 2000), whereas other work has suggested that this may produce a benefit (Stoffer, 1993).

More recent research has quantified attentional resizing costs by changing the proportion of trials where the target appears at different levels (e.g., target present at the global level 80 per cent of the time and at the local level 20 per cent of the time, and vice versa in a different block). This is designed to set attentional breadth at the level at which the target most often appears, and then the time to cost of resizing attention to the other level for the minority of trials can be gauged (Calcott & Berkman, 2014; Goodhew & Plummer, 2019). Using this method, average resizing costs around 100–150 ms have been observed, although with marked individual differences, such that resizing takes some individuals almost no time at all, and some individuals take up to 400–500 ms (Goodhew & Plummer, 2019) which is in the ballpark of the duration of other noteworthy deficits such as the attentional blink (Dell’Acqua et al., 2015; Raymond, Shapiro, & Arnell, 1992).

One suggestion has been that Navon stimuli induce changes in ‘categorical’ attention, rather than attentional breadth per se. That is, when participants attend to the global versus local elements of the stimuli, rather than this being mediated by changes in attentional breadth, participants may be essentially adopting attentional sets for *small, local elements* versus *large, global elements*, with their attention spread over equivalent spatial extents in both cases (Robertson et al., 1993). For a similar notion, see Coren, Ward, and Enns (2004). This is an interesting possibility to consider. However, one reason to doubt such a claim is that attending to the global versus local elements of Navon stimuli have been found to induce different attentional breadths, as revealed by functional magnetic resonance imaging (fMRI). Sasaki et al. (2001) used fMRI to measure activation in occipital areas in response to global and local (Navon) stimuli. Participants were instructed to either attend to or passively view hierarchical stimuli, and in the attend condition, participants were required to perform a shape identification task. Accuracy in this task was high. Sasaki et al. (2001) found that attending to the global stimuli activated more peripheral regions than for local stimuli and that the magnitude of activation for the local attention condition over the same area was greater than for the global attention condition, which is consistent with the concept of a zoom lens. This was definitely true in a number of extrastriate areas but was also true in the primary visual cortex (V1) for some participants. This is consistent with the zoom-lens model of attention (discussed in Section 3). More broadly, this is consistent with the notion of attentional breadth being implicated in identifying local versus global Navon targets, such that the spatial extent of the area over which attentional resources are spread changes according to whether participants attend to the global or local levels. This is good convergent evidence that behavioural-based Navon effects (e.g., faster responses at the global level) do likely reflect the adoption of a corresponding (e.g., broad) attentional breadth.

2.2 Kimchi and Palmer

Kimchi and Palmer stimuli have also been proposed to operationalise attentional breadth (Basso, Schefft, Ris, & Dember, 1996; Behrmann et al., 2006; Fredrickson & Branigan, 2005; Gasper & Clore, 2002; Kimchi & Palmer, 1982; Koldewyn, Jiang, Weigelt, & Kanwisher, 2013; Kramer, Ellenberg, Leonard, & Share, 1996; Pletzer, Scheuringer, & Scherndl, 2017). Typically, these stimuli are used in a task whereby participants are presented with a standard shape configuration (e.g., four smaller squares arranged in the shape of a square), and then with two other comparison configurations to choose from with respect to which the participant considers the most similar to the standard shape (see Figure 2). Of the two comparison options, one is consistent with the standard shape at the global level (e.g., a global square but local triangles), while the other is consistent with the standard shape at the local level (e.g., local squares but global triangle). This means that which of the two available options a participant chooses is thought to indicate whether their attention was directed to the global or the local level of the test stimulus and, therefore, whether they had a broad or narrow attentional breadth.

One potential downside with using the Kimchi and Palmer stimuli in this way is that given the nature of the task, it is typically only practical to ask participants to make such judgements about a relatively small number of stimulus configurations (e.g., 1–8). This means that the resulting measure is relatively coarse (e.g., a scale of 1–8 for how often they selected the globally-similar stimulus). However, even with this, performance on tasks using Kimchi and Palmer's stimuli are reliably associated with meaningful variance in other psychological processes (Basso et al., 1996; Dale & Arnell, 2015; Koldewyn et al., 2013). Performance-based measures of visual search have also been developed using these stimuli (Enns & Kingstone, 1995), which circumvent such issues.

Cambridge University Press
978-1-108-79621-7 — The Breadth of Visual Attention
Stephanie C. Goodhew
Excerpt
More Information

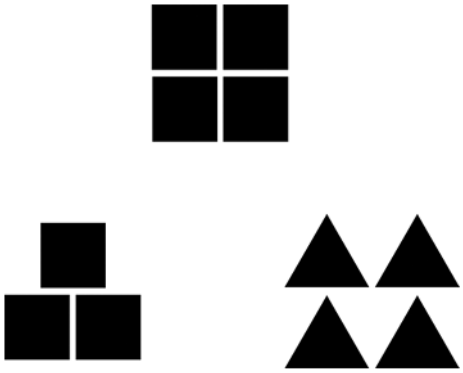

Figure 2 An example illustration of Kimchi and Palmer type stimuli. The top stimulus is a global square made up of local level squares. The bottom left stimulus (a triangle made up of squares) shares the same local elements as the top stimulus, but has a conflicting global configuration, whereas the bottom right stimulus (a square made up of triangles) has the same global configuration as the top stimulus, but conflicting local elements. Participants could be asked to indicate which of the two bottom stimuli they think looks most like the top stimulus.

2.3 Flanker

The flanker task is where a stimulus is presented, with surrounding task-irrelevant flanker stimuli (see Figure 3). For example, in the centre of the screen, the target letter ‘E’ could be presented, and the participants’ task is to identify whether the letter ‘E’ or ‘F’ is presented. One letter could appear to the left of the target and one letter could appear to the right of the target. These two flanking letters would be the same letter as one another, but crucially, they are either congruent (E E E) or incongruent (F E F) with respect to the target[2]. The flanker effect is defined as a slower response to the target when the flankers are incongruent, relative to when they are congruent (Biggs & Gibson, 2018; Eriksen & Eriksen, 1974; Richard, Lee, & Vecera, 2008). The spatial separation between the target and flankers can be varied, and the flanker effect typically diminishes as this separation increases (Eriksen & Eriksen, 1974).

The flanker effect is often classified as a measure of executive control of attention (Fan, McCandliss, Fossella, Flombaum, & Posner, 2005; Hotton,

---

[2] Note that congruency does not necessarily imply that the target and flankers have to be identical. For example, they can be not identical but be associated with the same (congruent) or different (incongruent) response. However, in contemporary work, this is one of the most common ways that congruency is operationalised: same identity flankers as the target (congruent) versus different identity flankers compared with the target (incongruent).