Episodic Visual Cognition
Implications for Object and Short-Term Recognition

Cumulative Dissertation

Submitted to the Faculty of Psychology and Sports Sciences,
Department of Psychology, at Bielefeld University
for the academic degree doctor rerum naturalium (Dr. rer. nat)

Submitted by Christian Hendrik Poth
Bielefeld, January 16, 2017

Contents

Acknowledgments

Summary

A great part of human behavior is based on visual cognition, the processing of visual information about external objects. For goal-directed behavior, two functions of visual cognition seem especially important. The first one is object recognition. Objects in the environment must be identified as belonging to an object category, so that they can be used to accomplish a given task. The second function is short-term recognition. It must be recognized whether an object in the environment has been viewed recently, so that current behavior involving the object can be related to previous behavior. Both functions share a common constraint: They must be fulfilled across distinct episodes of visual processing, which are interrupted by changes in processing demands.

For object recognition, visual processing episodes lead to a problem of *selective integration*. That is, it must be decided whether object information from the current episode should update and thus be integrated with object representations from the previous episode. Alternatively, object representations from two successive episodes are retained separately. This decision is critical. Updating and integration should enable a cumulative and fast object recognition. However, integration should also conceal object changes across episodes by leaving no separate representations that can be compared. Separation should improve change perception but impair object recognition, because limited visual processing resources for object recognition must be split between the separate representations.

For short-term recognition visual processing episodes lead to a problem of *matching*. That is, an object from the current episode must be matched against object representations, not only from the previous but from several episodes in the recent past (irrespective of whether objects are categorized).

The overarching goal of the present dissertation is to make a first step in understanding how the mechanisms underlying object recognition and short-term recognition operate across visual processing episodes, and how they solve the two problems. In five empirical studies, we investigated key issues that must be addressed before a theoretical account of object and short-term recognition across visual processing episodes can be given.

The first three studies focused on object recognition across visual processing episodes of eye fixations. Fixations are periods of visual information uptake, in which the eyes stand relatively still. They are separated by rapid saccadic eye movements. Saccades are necessary for object recognition, because they direct the central fovea of the eye's retina at interesting objects, allowing high-acuity inspection. However, saccades also disrupt visual input and displace and alter the retinal images of objects. Therefore, saccades dissect visual information processing into distinct episodes of fixations, which the mechanisms for object recognition must accommodate. In two studies (Poth, Herwig, & Schneider, 2015; Poth & Schneider, 2016a), we investigated how the selective integration problem is solved to support object recognition across successive fixations. We assessed the recent hypothesis (Schneider, 2013) that the problem is solved by a mechanism testing for correspondence ("object continuity") between an object before and after a saccade. If object correspondence is established, the object before and after the sac-

cade should be integrated into a common representation. In contrast, if object correspondence is broken, the object before and after the saccade should be represented separately. Separation should allow to compare the two representations, improving the discrimination of transsaccadic object displacements. At the same time, however, object recognition of the object after the saccade should be impaired, because the necessary visual processing resources had to be split between the two representations. Results were consistent with this hypothesis. Breaking object correspondence by briefly blanking an object after a saccade to it improved the discrimination of displacements of the object but impaired object recognition. Thus, the object correspondence mechanism seems to impact on object recognition after the saccade. Further experiments investigated the nature of object correspondence. They showed that object recognition was impaired when object correspondence was broken by changing an object's contrast-polarity (and luminance), its color-and-luminance, and its color alone. Together with the initial finding, this indicates that object correspondence is based on spatiotemporal as well as on the surface features of objects. In the third study (Poth & Schneider, 2016b, submitted), we went on to test the limits of object recognition across saccades. Because object recognition relies on limited visual processing resources, it can only be achieved for a few objects at a time. Here, we examined if different objects must compete for these resources across saccades. If this was the case, visual processing after a saccade would be slowed down as more and more objects are viewed before the saccade. Our findings show that this is the case, but only if the objects are task-relevant. Therefore, the findings support a key prediction of a recent theory, namely that the importance of an object representation determines whether it will survive a saccade and take up limited processing resources afterwards (Schneider, 2013).

With the fourth study (Poth & Schneider, 2016c), we turned from the processing episodes of successive eye fixations to those defined by appearing and disappearing objects and associated task-requirements. We asked about the relationship between the mechanisms underlying object recognition and those underlying short-term recognition. Visual processing for object recognition is assumed to be complete when an object has entered a limited-capacity visual working memory, where the object becomes available for being reported. We investigated if encoding into visual working memory is not only required for object recognition in the current episode, but also for short-term recognition in upcoming episodes. Supporting this notion, we found that objects that supposedly had not reached visual working memory were not available for later short-term recognition. This finding argues that the initial steps of visual processing before encoding into visual working memory are not sufficient for short-term recognition in later episodes. Therefore, visual working memory may contribute to the solution of the matching problem by limiting the amount of information considered in a short-term recognition task.

Finally, in the fifth study (Poth & Schneider, 2016d, submitted), we investigated short-term recognition further, asking how short-term recognition in a later processing episode can be prepared in advance. We assessed how prioritizing among objects represented in visual working memory impacts on two distinct components of performance in an upcoming short-term recognition task. Our results showed that such a prioritization improves memory-retention in visual working memory but also accelerates visual processing of objects for short-term recognition in a future episode. This indicates that changes in processing priorities contribute to ongoing solutions of the matching problem of short-term recognition.

Taken together, the five studies show how mechanisms of object and short-term recognition address specific problems arising from the dissection of visual processing into distinct episodes. As such, the studies implicate visual processing episodes as a source of problems for object and short-term recognition, which is neglected in most contemporary research. Conversely, however, the studies also invite speculation about the functional value of visual processing episodes for visual cognition.

Dissertation synopsis

Chapter 1

Introduction

1.1 Visual cognition: A cornerstone of human goal-directed behavior

Any goal-directed and intelligent behavior requires information about the environment it is situated in. For humans, a great part of this information is acquired visually, it is extracted from the light registered by the eyes (e.g., Marr, 1982; Palmer, 1999). Using this information for controlling action requires a great deal of further processing. An important part of this processing consists in visual cognition, creating and manipulating representations of external surfaces and objects (Cavanagh, 2011) from lower-level visual input (that is provided by the lower levels of the brain's visual hierarchy, for overviews, see Gilbert, 2013a; Meister & Tessier-Lavigne, 2013; Gilbert, 2013b; Albright, 2013).

Once processed up to a certain level, the object representations established by visual cognition can be used to perform goal-directed actions and they become accessible for report (e.g., Bundesen, 1990; Cavanagh, 2011; Schneider, 1995, 2013). One may say that the object representations are in a state of "access consciousness" (Block, 1995, 2011, cf. Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006). In this way, visual cognition plays an essential role in most human goal-directed behavior.

1.2 Functions and mechanisms of visual cognition

Visual cognition serves as a guide to human action. How visual cognition fulfills this purpose can be understood at different levels of explanation (Marr, 1982). For the present considerations, it is important to distinguish the *functions* of visual cognition from its *mechanisms*. The functions of visual cognition can be thought of as the goals of visual processing or computation, with respect to a given task or given action requirements (cf. Neumann, 1987, 1990). This has been called the "computational" level of explanation, because it dictates the overall strategy of visual processing (Marr, 1982).

How the functions of visual cognition are provided can be explained in terms of the mechanisms of visual cognition. Mechanisms specify input representations that are processed (cf. Palmer, 1978), output representations that result from processing, and the transformation converting input into output (the actual process). Defined in this way, the mechanisms are at what has been called the "algorithmic" level of explanation (Marr, 1982).

1.3 Two central functions of visual cognition: Object recognition and short-term recognition

Two functions of visual cognition are central components of most tasks humans perform to achieve their behavioral goals. The first function is object recognition: identifying external objects as belonging to a certain category and having certain features (Bundesen, 1990). Object recognition answers the question of which objects are there in the environment, which is a necessary requirement of using the objects to perform a task. The second function is short-term recognition: recognizing whether ob-

jects have been viewed recently (e.g., Kahana & Sekuler, 2002; Zhou, Kahana, & Sekuler, 2004; Poth & Schneider, 2016c). Short-term recognition answers the question of whether an object is encountered firstly in recent time or whether it has occurred before. This function is implied in all tasks in which specific object occurrences have to be tracked or discriminated over time.

1.4 Distinct visual processing episodes in object and short-term recognition

Object recognition and short-term recognition are part of a great deal of human goal-directed behavior, which is organized into tasks and subtasks of distinct task steps (e.g., Duncan, 2013; Land & Tatler, 2009; Norman & Shallice, 1986). While being engaged in a task, new processing demands can arise from changes in the environment and from the next task step (Schneider, 2013; cf. Duncan, 2013). This leads to a fundamental constraint of both of the two functions of visual cognition: They have to be accomplished across distinct visual processing episodes, which are characterized by specific processing demands and processing settings (Schneider, 2013).

The concept of the visual processing episode has been introduced by Schneider's (2013) theory of "Task-dRiven visual Attention and working Memory" (TRAM; note that visual processing episodes are called "competition episodes" there). In this theory, a new visual processing episode arises whenever the visual input or the processing demands of the current task change, and processing must be adapted accordingly.

In human vision, one ubiquitous type of visual processing episode consists in a fixation of relatively stable eye position, which is separated from the next by a rapid saccadic eye movement (for reviews, see Gegenfurtner, 2016; Rolfs, 2015; Schütz, Braun, & Gegenfurtner, 2011). Saccades are crucial for object recognition because they shift the central foveal region of the eye's retina toward potentially important objects, so that these are viewed with the highest visual acuity (cf. Stras-

burger, Rentschler, & Juttner, 2011). During saccades, however, vision is blurred and information uptake is suppressed (Krock & Moore, 2014; Wurtz, 2008). Sampling of visual information is therefore largely restricted to the intervals of eye fixations. In addition, each saccade drastically changes visual input. It changes the location of objects on the eye's retina, and due to the inhomogeneous visual resolution of the retina, this leads to changes in the visual acuity with which the objects are sampled (Curcio & Allen, 1990; Land & Tatler, 2009; Strasburger et al., 2011). Therefore, eye fixations constitute visual processing episodes, which are bounded by saccades, and which are distinct from one another due to the saccade-induced input changes (Schneider, 2013).

According to TRAM theory (Schneider, 2013), several factors create visual processing episodes in addition to saccadic eye movements. A new episode starts when objects appear or disappear in the visual field, or when object features change to a sufficient degree. Furthermore, visual processing episodes commence when visual processing priorities change, for instance when different objects become relevant for the current task or the next task step.

Taken together, there are a number of factors giving rise to distinct episodes in visual processing. It is an open question whether or not the visual processing episodes caused by each of the different factors are identical with respect to the mechanisms of visual cognition. However, irrespective of this question, it is clear that visual processing episodes pose challenges that these mechanisms must overcome.

1.5 Visual processing episodes as challenges for object and short-term recognition

The mechanisms underlying object recognition and short-term recognition operate over time and require time for processing. Visual processing episodes confine processing in time which leads to contrasting problems for these two kinds of mechanisms.

For the mechanisms underlying object recognition, visual processing episodes lead to what I call the *selective integration problem*. Successive processing episodes must be integrated in order to enable the cumulative acquisition of information about objects for their recognition (e.g., Demeyer, de Graef, Wagemans, & Verfaillie, 2009; Kahneman, Treisman, & Gibbs, 1992; Rayner, McConkie, & Zola, 1980). If this was not possible, then each new visual processing episode would force processing to start completely anew. This would impair or even prevent object recognition because new visual processing episodes may start so often that the time left for a single episode is less than needed for object recognition. For instance, visual processing episodes consisting in eye fixations are started anew by a saccade about every 250-300 ms (Land & Tatler, 2009) but object recognition can require that objects are viewed for longer durations (e.g., Petersen & Andersen, 2012; Shibuya & Bundesen, 1988). Likewise, new objects may appear or disappear and thereby start new processing episodes very often (Schneider, 2013), so that the duration of each episode falls short of what is necessary for object recognition. Furthermore, object recognition could be impaired because it relies on limited visual processing resources that had to be split among object representations if these were not integrated (Schneider, 2013). Thus, if there was no integration of processing across successive processing episodes, object recognition in active saccade-mediated vision, and in dynamic environments would be strongly hindered.

However, if object information was integrated across processing episodes in any circumstance, this would be likewise detrimental. Specifically, it would conceal changes and events in the environment that occur from one episode to the next (cf. Deubel, Schneider, & Bridgeman, 1996; Tas, Moore, & Hollingworth, 2012; Poth et al., 2015; Poth & Schneider, 2016a). For instance, object representations from the previous and the current episode could be integrated by updating or replacing the former with the latter (Schneider, 2013; cf.

Deubel et al., 1996), or by combining the two (e.g., Oostwoud Wijdenes, Marshall, & Bays, 2015; Wittenberg, Bremmer, & Wachtler, 2008). In both cases, changes of object features (e.g., color or location changes) across episodes would be undetectable, because no two representations were available for comparison (cf. Schneider, 2013). This would impair goal-directed behavior, because such changes of object features may be relevant to the current task or may otherwise signal important events in the environment (e.g., Rensink, 2002). Therefore, the selective integration problem is, at its core, that it must be decided whether object information should be integrated or separated across visual processing episodes.

For the mechanisms underlying object recognition, the current visual processing episode must be reconciled with the previous one to establish the object's features or object category. For the mechanisms underlying short-term recognition, visual processing episodes pose a different problem, one of *matching*. Short-term recognition means determining whether a currently present object has been viewed recently. Hence, an object from the current visual processing episode must be matched against the objects of multiple recent episodes, irrespective of whether objects are categorized (as in object recognition). Such a matching requires to distinguish the content of visual processing episodes, especially if similar objects can appear in multiple visual processing episodes. Therefore, the function of short-term recognition is intrinsically linked to visual processing episodes. For this reason, we also called the function "episodic short-term recognition" (Poth & Schneider, 2016c).

To fulfill their functions, the mechanisms underlying object recognition and short-term recognition must accommodate visual processing episodes. It is unclear, however, how this is accomplished, how the mechanisms enable processing across episodes, and whether and how they work in concert to this end.

4

1.6 The present dissertation

The present dissertation aims at shedding light on how the mechanisms underlying object recognition and short-term recognition operate across visual processing episodes. Five empirical studies investigated issues that must be addressed as a first step toward a theoretical account of such visual cognition across visual processing episodes. As such, this dissertation is meant to be a starting point and a call for research on the topic of episodic visual cognition. The following chapters of the dissertation synopsis discuss the theoretical background and the findings of the empirical studies. The original studies are provided afterwards, as the final part of the dissertation.

Chapter 2 provides the theoretical background for studying how the mechanisms underlying object recognition accommodate the challenges of visual processing episodes. To this end, the part introduces theories of visual attention (Bundesen, 1990; Bundesen, Habekost, & Kyllingsbæk, 2005; Desimone & Duncan, 1995) and visual working memory (VWM; e.g., Bundesen, Habekost, & Kyllingsbæk, 2011; Eriksson, Vogel, Lansner, Bergstrom, & Nyberg, 2015; Luck & Vogel, 2013; Olivers, Peters, Houtkamp, & Roelfsema, 2011), two heavily intertwined mechanisms that together accomplish object recognition within a single processing episode. On this basis, Schneider's (2013) TRAM theory is presented, which extends the mechanisms to processing across episodes.

Building on the theoretical background, Chapter 3 discusses three empirical studies that investigated these mechanisms for the visual processing episodes consisting in eye fixations, which are separated by saccadic eye movements. In Study 1 (Poth et al., 2015) and Study 2 (Poth & Schneider, 2016a), we assessed a central proposal of TRAM theory (Schneider, 2013), namely, that there is a mechanism establishing correspondence between representations of an external object before and after the intervening saccade, which impacts on subsequent object recognition. In this way, the two studies investigated TRAM's solution to the selective integration problem of object recognition across visual processing episodes. In Study 3 (Poth & Schneider, 2016b, submitted), we went on to study the limits of object recognition across saccades. Object recognition can only be achieved for a few objects at a time (for reviews, see Bundesen & Habekost, 2008; Duncan, 2006). Therefore, we examined how this limited capacity is distributed across successive fixations for object recognition.

Chapter 4 asks about the relationship between the mechanisms underlying object recognition and those underlying short-term recognition. To address this issue, in Study 4 (Poth & Schneider, 2016c) we investigated whether visual processing for recognition of an object must have been completed for short-term recognition of the object in a later processing episode. In Study 5 (Poth & Schneider, 2016d, submitted) we then turned to the question by what means such a short-term recognition in a later processing episode can be prepared in advance. To this end, we assessed how prioritizing among object representations retained from a previous episode impacts on two distinct components of performance in an upcoming short-term recognition task.

Finally, chapter 5 aims to offer a brief general discussion of how the five studies contribute to our understanding of object recognition and short-term recognition across visual processing episodes. Up to this point, visual processing episodes have been regarded as a processing requirement and a challenge. Abstracting from the mechanisms of object and short-term recognition addressing this challenge, Chapter 5 also presents some speculations about the functional value of visual processing episodes for visual cognition.

Chapter 2

Theoretical background: Mechanisms for visual cognition within and across visual processing episodes

2.1 Visual attention and working memory

Human capacity for object recognition is limited: not all objects in the visual field can be recognized at the same time (for reviews, see Bundesen & Habekost, 2008; Duncan, 2006). As proposed by the *biased competition* framework (Desimone & Duncan, 1995), the objects compete against each other for object recognition (Desimone & Duncan, 1995). Object recognition for task-driven behavior therefore requires to select currently relevant objects for being recognized, at the expense of irrelevant ones. This function is called *selection-for-perception* (Schneider, 1995; Schneider & Deubel, 2002).

Human capacity for acting upon objects is limited as well, because there are only a few effectors (e.g., two hands) that can be used for this purpose (Neumann, 1987, see also Allport, 1987; Neumann, 1990). Again, the selection of relevant over irrelevant objects is necessary. This function is called *selection-for-action* (Allport, 1987). Selection-for-perception and selection-for-action are assumed to be performed by common mechanisms of *visual attention* (Deubel & Schneider, 1996; Schneider, 1995; Schneider & Deubel, 2002).

A theory of how the mechanisms of visual attention mediate object recognition within a single processing episode has been provided by Bundesen's (1990) "Theory of Visual Attention" (TVA; for more recent reviews, see Bundesen &

Habekost, 2008; Bundesen, Vangkilde, & Petersen, 2015). This theory provides a widely-used framework for studying visual attention, not only in experimental psychology (for reviews, see Bundesen & Habekost, 2008; Bundesen et al., 2015), but also neuropsychology (e.g., Duncan et al., 1999; Finke et al., 2005; Finke, Bublak, Dose, Müller, & Schneider, 2006; reviewed by Habekost, 2015), and clinical diagnostics (e.g., Habekost, Petersen, & Vangkilde, 2014; Foerster, Poth, Behler, Botsch, & Schneider, 2016).

TVA (Bundesen, 1990; Bundesen et al., 2005) assumes that external objects are recognized, when the objects become represented in VWM (in TVA called "visual short-term memory"). VWM retains information about a limited number of objects over short time-windows, so that this information can be reported (Luck & Vogel, 1997, 2013; Eriksson et al., 2015; Shibuya & Bundesen, 1988; but see Bays, 2015; Ma, Husain, & Bays, 2014) or otherwise used for action (e.g., Schneider, 2013).

Whether or not objects become represented in VWM depends on visual processing. That is, in TVA (Bundesen, 1990) objects enter VWM if there is sufficient retention space and if visual processing of the objects has been completed. Visual processing is assumed to proceed in two stages. In the first stage, visual input is segmented into perceptual units corresponding to external objects. The information provided by these perceptual units is subsequently compared with visual features and

categories that have been acquired with experience and reside in visual long-term memory. This comparison yields values of sensory evidence that an object has a certain feature or belongs to a certain category. Importantly, the comparison proceeds for all objects in the visual field and is thus unselective.

In contrast to the first stage, processing at the second stage is selective. The categorizations of objects in the visual field are assumed to participate in a competitive race toward VWM. The object categorizations that finish processing first are encoded into VWM (if enough retention space is available there) and this allows all other categorizations of the same objects to become represented there as well. Encoding into VWM only continues until it is filled up with categorizations of a limited number of different objects. On the neuronal level (Bundesen et al., 2005), VWM is assumed to consist in a topographically organized map representing objects, henceforth called the *VWM map of objects* (following Schneider, 2013), and loops of neuronal activity between these objects and neurons coding for the visual features belonging to the objects (Bundesen et al., 2005). By means of these loops, the activity of the neurons representing visual object features is sustained and can outlast the presence of the external object in the visual field.

Two mechanisms of visual attention jointly determine the speed with which the categorization of an object is processed, that is, the speed with which the categorization races towards VWM. The first is the *pigeonholing* mechanism. This mechanism consists in an internal perceptual decision bias for categorizing any object as having a specific feature. That is, the bias multiplicatively weights the sensory evidence for task-relevant features, so that irrespective of which objects are actually viewed, categorizations of this feature are processed faster. Neuronally, the pigeonholing may be implemented as an increased firing rate of all those neurons preferentially coding for the feature in question (Bundesen et al., 2005).

For a given object, the speed with which its categorizations are processed is proportional to the amount of visual processing resources allocated to the object. This is where the second mechanism comes into play, which is called *filtering*. This mechanism operates by assigning each object in the visual field an attentional weight. The attentional weight reflects the current importance of the object. The weight is computed at the first stage of processing by summing up the sensory evidences that the object has certain features, whereby the evidence for a feature is multiplicatively weighted by the current importance of the feature. In this fashion, the attentional weight combines bottom-up information consisting in sensory evidence for a feature with top-down information consisting in the importance of this feature (additional bottom-up factors of attentional selection may also contribute to filtering, Nordfang, Dyrholm, & Bundesen, 2013, and attentional weights can be used to monitor for object changes with low expected bottom-up salience, Poth, Petersen, Bundesen, & Schneider, 2014). Visual processing resources are allocated to an object according to the object's attentional weight relative to the sum of the attentional weights of all objects in the visual field. As a result, the categorizations of currently important objects are processed fastest. In this way, the objects are selected for object recognition.

According to the neural interpretation of TVA (NTVA; Bundesen et al., 2005), the attentional weights are stored in a spatially organized priority map (see also Bundesen et al., 2011). Priority maps exist in several areas of the primate brain, such as monkeys' frontal eye field (Schall, 2009), lateral intraparietal area (Bisley & Goldberg, 2010), pulvinar (Kastner & Pinsk, 2004), and superior colliculus (Krauzlis, Lovejoy, & Zenon, 2013). In general, they combine the bottom-up salience (intrinsic to external objects) with the top-down task-relevance of objects or features (Fecteau & Munoz, 2006; Zelinsky & Bisley, 2015).

In NTVA, the attentional weights of the priority map set gates within the visual system's ventral stream for object recognition. The receptive fields (the regions of the retina from which they receive

input) of higher level neurons (e.g., in the inferior temporal cortex) are dynamically remapped so that they receive input from lower level neurons coding for visual features of a specific object. In this manner, more neurons are allocated to objects with high attentional weights than to those with lower ones. Thus, these neurons are the visual processing resources that are distributed across objects according to their attentional weights.

The filtering and pigeonholing mechanisms proposed in TVA (Bundesen, 1990) explain how visual attention selects objects and visual object features for object recognition. The selection in TVA is restricted to the situation of a single processing episode. This is necessary to develop a mechanistic account of the selection, but falls short of situations with more than one processing episode. This issue is addressed by Schneider's (2013) TRAM theory, which extends the mechanisms of visual attention (specifically the filtering mechanism) and of VWM to processing across visual processing episodes.

2.2 Visual attention and working memory across visual processing episodes

As explained above, TRAM theory (Schneider, 2013) assumes that visual processing for object recognition is structured in visual processing episodes ("competition episodes"). In TRAM, a single visual processing episode consists of three processing phases, whereby the first two phases correspond to the two processing stages of TVA (Bundesen, 1990; Bundesen et al., 2005).

The first phase is the unselective stage of TVA (Bundesen, 1990; Bundesen et al., 2005), in which attentional weights of objects in the visual field are computed. Extending TVA, however, TRAM (Schneider, 2013) assumes that the attentional weights in the priority map form part of so-called *proto-objects*, candidate object representations for object recognition (Wischnewski, Steil, Kehrer, & Schneider, 2009; Wischnewski, Belardinelli, Schneider, & Steil, 2010, cf. Rensink, 2000). Besides an attentional weight, a proto-object represents the rough location and shape of an external object, as these two features are coded by the priority map additionally. Furthermore, a proto-object comprises visual object features, which are represented in the ventral and dorsal streams of the brain's visual system (Wischnewski et al., 2009, 2010).

The second phase of TRAM (Schneider, 2013) corresponds to the competitive race towards VWM in TVA (Bundesen, 1990; Bundesen et al., 2005). In TRAM, the proto-objects compete against each other for encoding of their features into VWM and for being recognized in this fashion. Proto-objects are inaccessible for being reported. However, when the features of a proto-object enter VWM, then the proto-object is transformed into an object representation in VWM. If this has happened, the represented object can be reported.

In the third phase of processing, those object representations in VWM that are task-relevant are made available for controlling behavior beyond the current processing episode. As in NTVA (Bundesen et al., 2005), retention in VWM is assumed to consist in sustaining activity in feature-coding neurons by looping neuronal activity between them and the VWM map of objects (Bundesen et al., 2005). TRAM assumes that with ongoing retention in VWM, the retained object representations are consolidated into a *passive* state that does not require such a looping of activity any more (Larocque, Lewis-Peacock, & Postle, 2014; Stokes, 2015; and may rely on changes in synaptic connectivity, e.g., Mongillo, Barak, & Tsodyks, 2008; Rose et al., 2016). In this passive state, the represented information can be retained across the several visual processing episodes without taking up retention space in capacity-limited VWM. This is important because retained information would otherwise block encoding of new objects into VWM in the next episodes. Such a blocking would impair or even prevent the performance of tasks requiring visual guidance over multiple episodes, such as multi-step sensorimotor tasks (e.g., Hayhoe & Ballard, 2005; Foerster, Car-

bone, Koesling, & Schneider, 2011).

After TRAM's (Schneider, 2013) third phase of processing has finished, the next visual processing episode starts with its first phase. Importantly, visual processing episodes are linked by the interplay of VWM and the attentional weights that mediate encoding into VWM. More specifically, the task-relevant objects that are represented in VWM remain connected to their attentional weights (cf. Petersen, Kyllingsbæk, & Bundesen, 2012, 2013). Consequently, the attentional weights of these objects from the previous episode persist in the current episode. Both, the attentional weights and the object representations in VWM are linked to the representations of visual features. These links are critical for integrating information across visual processing episodes. For a given external object, the attentional weight in a new visual processing episode matches the one from the previous episode (or a prediction of this attentional weight, respectively). New visual input that arrives at the representations of visual features can therefore be routed by the attentional weight to the respective object in VWM. As a result, the object representation in VWM is *updated* by new visual information.

Importantly, if this updating refers to a feature that has already been represented as part of the object in VWM, then the representation of this feature is overwritten (or combined) with the new information about it (Schneider, 2013; cf. Poth et al., 2015; Poth & Schneider, 2016a). In contrast, if the updating refers to a feature that has not been part of the object in VWM, then this feature can be attached to the object in VWM additionally (the VWM object is said to be "re-categorized" with a new feature, Schneider, 2013, p. 8-9).

The described updating mechanism provides a means for integrating information about external objects sampled in successive processing episodes. Importantly, however, this does not yet solve the selective integration problem of object recognition across visual processing episodes. At the heart of this problem is the decision whether integration should take place or whether objects from suc-

cessive episodes should be represented separately. As solution to this problem, TRAM (Schneider, 2013) proposes a mechanism that tests for correspondence ("object continuity") between objects in VWM from the previous episode and the objects of the current episode. This test consists in a comparison of the objects retained in VWM with objects of the current episode in terms of their features in a priority map, namely their attentional weights, locations, and rough shapes (or more precisely, predictions derived from these priority map features).

If the test for object correspondence is positive for an object, then this implies that the object is processed with an identical attentional weight across the visual processing episodes. In this case, the attentional weight is used to update the representation of the object in VWM with new visual input. In contrast, if the test for object correspondence is negative (i.e. object correspondence is broken), then the attentional weight of the object from the previous episode is *encapsulated*, meaning that it is retained in its current state. This shields the associated object in VWM from being updated with new visual input. Thus, new visual input is then treated as belonging to a new external object and has to pass through all processing phases (even if the input actually stems from the same external object that resulted in the VWM representation with the encapsulated attentional weight).

Encapsulating an object's attentional weight has attentional costs. The neuronal processing resources belonging to the attentional weights cannot be used for processing other, subsequent objects (see Petersen et al., 2012 for related ideas). Fewer resources are thus available for processing new objects, so that these are processed more slowly and their object recognition is impaired.

Critically, the encapsulation of attentional weights is assumed to happen only for task-relevant objects in VWM. Furthermore, the encapsulation happens only as long as these objects are in the third phase of processing which ends when they have been consolidated into the passive state

of VWM. Thus, once the objects are in the passive state, they can be made available for action control (by retrieval into VWM) without permanently requiring retention space in VWM and without causing attentional costs for object recognition of subsequent objects.

In sum, two of TRAM's proposals are essential for object recognition across visual processing episodes. First, the proposal that there is a mechanism that tests for object correspondence across visual processing episodes. This mechanism impacts on subsequent object recognition by deciding whether attentional weights are encapsulated or used for updating. As such, this mechanism offers a hypothetical solution to the selective integration problem for object recognition across visual processing episodes. Second, the proposal that only those attentional weights are encapsulated that belong to task-relevant objects in VWM. This proposal implies that attentional competition across episodes is limited to task-relevant objects, preventing that object recognition in general suffers with each new visual processing episode.

Chapter 3

Object recognition across visual processing episodes

3.1 Object correspondence linking the visual processing episodes of successive eye fixations

TRAM theory (Schneider, 2013) presents a functional view on object correspondence, assuming that it serves object recognition across visual processing episodes (Schneider, 2013). Traditionally, however, object correspondence mechanisms have mostly been held to explain phenomena of visual perception (which may also support behavior by carrying information about regularities in the world, e.g., Flombaum, Scholl, & Santos, 2009). Among the most prominent examples is the percept that one and the same object is present at successive locations in apparent motion (e.g., Kahneman et al., 1992; Mitroff & Alvarez, 2007) and across occlusion by another object (e.g., Hollingworth & Franconeri, 2009). Most importantly to visual processing across episodes, object correspondence has recently been taken to explain visual stability across the episodes of eye fixations separated by saccadic eye movements (Tas et al., 2012; Tas, 2015; cf. Hollingworth, Richard, & Luck, 2008).

Visual stability is the phenomenon that external objects are perceived as stable in their locations across a saccade, despite the fact that the saccade changes their locations on the eye's retina and their visual resolution (for reviews, see Higgins & Rayner, 2015; Mathôt & Theeuwes, 2011; Rolfs, 2015; Wurtz, 2008). Interestingly, the perception of visual stability arises even despite actual object displacements: Displacing an object while a saccade is made towards it is hard to notice, even for relatively large displacements (of up to about a third of saccade amplitude, Bridgeman, Hendry, & Stark, 1975, on the saccade axis, Wexler & Collins, 2014). The extent of this insensitivity to transsaccadic displacements provides an often used measure of visual stability in laboratory experiments (e.g., Deubel et al., 1996; Deubel, Bridgeman, & Schneider, 1998; for a review, see Higgins & Rayner, 2015)

For a long time, it has been assumed that the inability to detect (or discriminate) transsaccadic displacements stems from a reset of visual processing after a saccade, which leads to the loss of presaccadic location information (Bridgeman, van der Heijden, & Velichkovsky, 1994). This hypothesis was, however, falsified with the *blanking effect* (Deubel & Schneider, 1994; Deubel et al., 1996, 1998). That is, the insensitivity to transsaccadic displacements is ameliorated when the object is briefly blanked after the saccade and then reappears at its displaced location. Under such conditions, the direction of the displacement can be reported accurately. This finding reveals that there is information about the presaccadic object location after the saccade, but that this information is usually inaccessible, in the absence of the blank.

In the light of TRAM (Schneider, 2013), the blanking effect may be interpreted as follows (see also Poth et al., 2015; Poth & Schneider, 2016a). After eye-landing, the visual system tests for correspondence ("object continuity") between the presaccadic object and the postsaccadic object (the

discussion in terms of *object correspondence* was initiated by Tas et al., 2012). If the test for transsaccadic object correspondence is positive, the VWM representation of the presaccadic object is updated after the saccade with input from the postsaccadic object. As a result, only a combined representation of the presaccadic and postsaccadic object is available in VWM. This prevents any comparison between the presaccadic and postsaccadic location. Transsaccadic location changes (and visual feature changes, Deubel, Schneider, & Bridgeman, 2002; Weiß, Schneider, & Herwig, 2015) are therefore indiscriminable and visual stability is perceived.

In contrast, if the eye lands on an empty display because the object is blanked, then the test for transsaccadic object correspondence is negative (because there is a mismatch between the attentional weight of the object from before the saccade and from the blank display after the saccade, see also section 2.2). Consequently, the attentional weight of the presaccadic object that is now in VWM is encapsulated. The object representation is therefore maintained as is, rather than updated with new input. A new representation is then created for the postsaccadic object after the blank. The two distinct representations can be compared, which improves the discrimination of transsaccadic displacements and prevents the perception of visual stability. However, encapsulating the attentional weight of the presaccadic object representation means that the representation takes up visual processing resources required for processing the postsaccadic object. Therefore, TRAM predicts that it should impair recognition of the postsaccadic object.

The goal of **Study 1** (Poth et al., 2015) was to test the hypothesis that breaking object correspondence across the saccade improves perception of transsaccadic displacements but at the same time impairs postsaccadic object recognition. For this purpose, Experiment 1 of Study 1 combined the classic blanking paradigm (Deubel & Schneider, 1994; Deubel et al., 1996) with an object recognition task. Participants made saccades to a peripheral saccade target object containing an irrelevant special character. During the saccade, the object was displaced, and after eye-landing a letter was shown in the object and terminated by a pattern mask. The postsaccadic object and the letter in it were visible either immediately after eye-landing, or after a brief blank display. Critically, in one block of trials, participants discriminated the direction of the object displacement. However, in another block of trials, they reported the identity of the postsaccadic letter.

Replicating the blanking effect, participants' displacement discrimination was more accurate when there was a postsaccadic blank compared with when the object was shown immediately after the saccade. Thus, breaking object correspondence by blanking improved the perception of transsaccadic displacements and reduced perceived visual stability. In contrast, however, participants' letter reports were less accurate when there was a postsaccadic blank than when there was none. Thus, breaking object correspondence by blanking impaired postsaccadic object recognition. Experiment 2 followed up on this latter result to rule out a number of alternative explanations based on the object displacement and the delayed onset of the postsaccadic object in case of a blank. In this experiment, participants only reported the letter that was shown in the postsaccadic object, and there were no object displacements across the saccade. In contrast to Experiment 1, object correspondence was broken by changing the contrast-polarity (and luminance) of the object across the saccade (this manipulation followed Tas et al., 2012). A black object on a gray background was changed into a white one during the saccade and vice versa. Participants' letter report performance was compared between conditions with and without such contrast-polarity changes. Thus, here the time-course of object appearances was identical in all experimental conditions, because there was no blank delaying the onset of the postsaccadic object. Despite the different manipulation of object correspondence across the saccade, the results of Experiment 2 were consistent with those of Experiment

1. Recognition of the postsaccadic object was impaired when object correspondence was broken by the contrast-polarity change, compared with when no such change occurred. Thus, taken together, the results of Study 1 offer evidence that breaking object correspondence across saccadic eye movements impairs postsaccadic object recognition.

Study 1 manipulated transsaccadic object correspondence by means of blanking and changes of the contrast-polarity of achromatic objects. Both manipulations have strong perceptual effects (Deubel et al., 1996, 1998; Tas et al., 2012). This may imply especially strong effects on object correspondence, because of a substantial contribution to the priority map features on whose basis object correspondence is tested for (Schneider, 2013). One may therefore ask whether the findings of impaired object recognition generalize to other object features and more subtle feature changes.

One of the most important surface features for human object recognition is color (Gegenfurtner, 2003; Gegenfurtner & Kiper, 2003; Moutoussis, 2015). However, testing object correspondence across saccades based on color may be problematic. Specifically, the color of an object is represented with strikingly different quality and resolution at the fovea of the retina, corresponding to gaze center, compared with the retinal periphery (e.g., Hibino, 1992; Johnson, 1986; Livingstone & Hubel, 1987; Nagy & Wolf, 1993). Every saccade displaces an object's image on the retina, so that the presaccadic and postsaccadic object image differ in color quality. Thus, every saccade induces an object change in terms of its color input. According to TRAM (Schneider, 2013) and based on our previous findings (Poth et al., 2015), if such naturally occurring color changes broke transsaccadic object correspondence, then postsaccadic object recognition would be impaired with every saccade. Therefore, one may hypothesize that color is ignored in the test for transsaccadic object correspondence in order to salvage postsaccadic object recognition.

The goal of **Study 2** (Poth & Schneider, 2016a) was therefore to test whether breaking object correspondence with color changes, and thus with more subtle surface feature changes, also impairs postsaccadic object recognition. Study 2 adopted the paradigm of the second experiment of Study 1 (Poth et al., 2015). In Experiment 1 of Study 2, transsaccadic object correspondence was broken by changing the object's color-and-luminance across the saccade, that is, by changing between red and green of different luminances. In Experiment 2, transsaccadic object correspondence was broken by changing the object's color alone, by changing between blue and yellow of about the same luminance. In both experiments, breaking object correspondence across the saccade impaired postsaccadic object recognition. These results show that the surface features of color-and-luminance as well as the one of color alone contribute to object correspondence and thereby impact on object recognition.

Taken together, the findings of Study 1 (Poth et al., 2015) and Study 2 (Poth & Schneider, 2016a) reveal that the mechanism establishing object correspondence across the saccade not only determines the perception of visual stability but also impacts on object recognition. Thus, the findings support the prediction of TRAM theory (Schneider, 2013) that object recognition across visual processing episodes, here successive eye fixations, depends on a mechanism that tests for object correspondence.

Such a mechanism for object correspondence presents one solution to the selective integration problem of object recognition across visual processing episodes. Specifically, the mechanism decides whether object information should be integrated across the saccade or whether presaccadic and postsaccadic object representations should be kept separate (cf. Deubel et al., 1996; Tas et al., 2012). The integration of object information across saccades may enhance object recognition, by preventing attentional competition between representations (i.e. their attentional weights, Schneider, 2013) and by enabling a cumulative visual processing (Demeyer et al., 2009; Henderson & Anes,

1994; Kahneman et al., 1992). In contrast, the separation of object representations may enhance comparisons across the saccade and thus help discriminate transsaccadic object changes (e.g., Deubel et al., 1996, 2002; Weiß et al., 2015).

3.1.1 Object correspondence for object recognition based on multiple object features

Besides revealing a link between object correspondence across the saccade and object recognition, the results of Study 1 and 2 also shed some light on the object correspondence mechanism itself. Specifically, they show that the mechanism takes multiple object features into account. The question which features are used to establish object correspondence is also hotly debated in the literature on object correspondence across occlusion (Hollingworth & Franconeri, 2009) and apparent motion (Kahneman et al., 1992; Mitroff & Alvarez, 2007). Some accounts maintain that object correspondence is based on spatiotemporal features only (Kahneman et al., 1992) or at least primarily (Flombaum et al., 2009; Scholl, 2007). This is consistent with the view that object correspondence across saccades is broken by blanking, because blanking implies a mismatch between presaccadic and postsaccadic objects in space and (expected) time.

However, we also found postsaccadic object recognition impaired when transsaccadic object correspondence was broken by changes of contrast-polarity (and luminance) of achromatic objects, changes of color-and-luminance, and of color alone. Thus, such surface features contribute to the test for transsaccadic object correspondence as well. Converging evidence for this proposal comes from two earlier studies. The first one showed that changing contrast-polarity and more complex visual features (pictures of real-world objects) improved the discrimination of transsaccadic displacements and thus diminished the perception of visual stability (Tas et al., 2012). The second study induced transsaccadic object displacements that caused a saccade to land between two objects

(Hollingworth et al., 2008). Corrective saccades in response to these displacements were guided by the objects' color-and-luminance, which offers indirect evidence that the features contribute to object correspondence.

As our findings, the ones of the two studies argue that surface features and not only spatiotemporal features contribute to transsaccadic object correspondence. A similar contribution of surface features has been observed for object correspondence across occlusion and motion as well (Hollingworth & Franconeri, 2009). Thus, we may now conclude that object correspondence across visual processing episodes generally takes spatiotemporal as well as surface features into account.

3.1.2 Object correspondence based on attentional weights and predictive remapping of receptive fields

How could a test for transsaccadic object correspondence be implemented? According to TRAM, transsaccadic object correspondence is tested for by comparing the attentional weight of the object (and other characteristics in the priority map) after the saccade with a prediction that has been derived from the presaccadic attentional weight (Schneider, 2013). This prevents that predictable changes of sensory input due to the saccade-induced shift of the object's retinal image impact on object correspondence (cf. Herwig & Schneider, 2014). For the spatial location of an object, this prediction may be implemented by *predictive remapping* (Duhamel, Colby, & Goldberg, 1992; for reviews, see Wurtz, 2008; Wurtz, Joiner, & Berman, 2011; but see, Zirnsak, Steinmetz, Noudoost, Xu, & Moore, 2014). Just before a saccade is made, neurons in some retinotopically organized brain areas become responsive to the locations their receptive fields will be brought to by the saccade. These brain areas include the ones assumed to implement priority maps (monkeys' frontal eye fields, Umeno & Goldberg, 1997; lateral intraparietal areas, Duhamel et al., 1992, and superior colliculi, Walker, Fitzgibbon, & Gold-

berg, 1995). The predictive remapping of receptive fields is based on a corollary discharge (efference copy) signal from saccade-generating neurons in the superior colliculus, which informs about the amplitude and direction of the impending saccade (Sommer & Wurtz, 2006). This may allow to track the approximate location of objects across saccades. Specifically, predictive remapping may make a given neuron receive input from one and the same object before and after a saccade (Wurtz et al., 2011) and this seems to contribute to the perception of visual stability (Cavanaugh, Berman, Joiner, & Wurtz, 2016; which should, however, tolerate transsaccadic object displacements to a degree, cf. Bridgeman et al., 1975).

However, predictive remapping of receptive fields is limited to object locations, it does not allow to track surface features of objects (Cavanagh, Hunt, Afraz, & Rolfs, 2010). A specification of TRAM theory's (Schneider, 2013) object correspondence mechanism may address this problem (see, also Poth & Schneider, 2016a). That is, predictive remapping may configure the attentional weights in a priority map for the impending object correspondence test after the saccade. According to NTVA (Bundesen et al., 2005), the attentional weights control the dynamic remapping of neuron's receptive fields in the higher ventral stream. The attentional weights set gates in the ventral stream, so that higher level neurons receive input from the lower level neurons coding for the visual features of a specific external object. Predictive remapping may now reconfigure the attentional weights, so that they already set the gates in accordance with the retinal locations objects will fall on after the saccade (see, Cavanagh et al., 2010 for a related idea of remapping of "attention pointers"). Then, after the saccade, the higher level neurons should receive feature input from the objects at the locations dictated by the attentional weights. This feature input may in turn retain the attentional weights, because it sets the sensory evidence values from which the attentional weights are computed (see the above description of TVA, Bundesen, 1990) in this fixation (Schneider,

2013). Thus, when the same object is present before and after the saccade, the attentional weight remains unchanged. In TRAM, the test for object correspondence is then positive and the feature input is routed through the gates set by the attentional weights to update the presaccadically created VWM representation of the object.

In contrast, when the object is blanked after the saccade or when its surface features change, then the feature input to the attentional weight is missing. This results in a negative test for object correspondence. As proposed by TRAM (Schneider, 2013), the attentional weight should then be encapsulated. As a consequence, the higher level neurons whose gates are set by the attentional weight cannot be redistributed to receive input from new objects. This cuts the resources for processing these objects which offers one explanation for the impaired object recognition that we observed in Study 1 and 2, when object correspondence was broken across the saccade (Poth et al., 2015; Poth & Schneider, 2016a). Besides this explanation, TRAM also provides another, not mutually exclusive explanation of this effect. Namely, object recognition could have been impaired because breaking object correspondence prevented the integration and updating of the representation of the object in VWM (for a discussion, see Poth et al., 2015; Poth & Schneider, 2016a). That is, the postsaccadic object would not have been processed by using the same attentional weight as the presaccadic one, preventing that the postsaccdic features are routed to the presaccadically created object representation in VWM. Consequently, a new attentional weight and a new representation in VWM would have to be created for the postsaccadic object. This would delay processing and should impair postsaccadic object recognition, especially if the object was terminated by a mask as in our present studies.

3.2 Attentional competition for object recognition across the visual processing episodes of successive eye fixations

Object correspondence may determine whether or not processing of an object is integrated across visual processing episodes. As we have seen so far, this should decide whether visual processing resources are allocated to one integrated object representation (using one attentional weight) or two separate object representations from each episode (using two attentional weights; Schneider, 2013). In most situations and visually-guided tasks, a number of objects are present in the visual field and may appear, disappear, or change from one processing episode to the next. Therefore, another question of fundamental importance is how processing resources for object recognition are distributed across different external objects in successive visual processing episodes.

Within a visual processing episode, such as an eye fixation, different objects must compete for the limited visual processing resources (Bundesen, 1990; Duncan, 2006, such as neurons, Bundesen et al., 2005). Mechanisms of visual attention bias this competition, so that currently important objects receive more resources than unimportant ones (Desimone & Duncan, 1995). The filtering mechanism of TVA (Bundesen, 1990; see section 2.1) is a specification of such a mechanism. As described above, it operates by distributing resources across objects according to attentional weights reflecting the objects' current importance. However, this mechanism does not allow a differential allocation of processing resources for objects of equal importance. Therefore, the more equally important objects enter the visual field, the smaller is the amount of processing resources allocated to each individual object, slowing down the objects' processing for object recognition (Bundesen, 1990).

Surprisingly, while a great deal of research has been devoted to understanding attentional competition for object recognition within eye fixations, it is unclear whether it also extends from one fixation to the next. In other words, it is unclear whether objects from one fixation compete with those in the next fixation for object recognition.

Three hypotheses can be advanced as answers to the question of whether objects compete for recognition from one fixation to the next. The first is that successive fixations are entirely separate episodes with no competition between the objects in them. This is in line with the proposal that to survive the saccade, objects have to be encoded into VWM (Irwin, 1992; Irwin & Gordon, 1998). The competition for object recognition is, however, assumed to rely on object representations created before encoding into VWM (Bundesen, 1990; Bundesen et al., 2005). Therefore, these object representations before VWM should not exist beyond the current fixation and there should be no attentional competition across the saccade per se.

However, some object information outside VWM seems to persist across the saccade, although it is largely bound to the retinal object locations which are changed by the saccade (Irwin, Brown, & Sun, 1988). This argues against the first hypothesis, because the competition for object recognition includes all objects in the visual field and should thus include these persisting object representations (Bundesen, 1990; Bundesen et al., 2005; Desimone & Duncan, 1995). Therefore, the second hypothesis predicts that these persisting object representations should compete with and impair processing of objects in the next fixation.

The third hypothesis directly follows from TRAM theory (Schneider, 2013), objects should compete for object recognition across saccades, but only if they are relevant to the task at hand. That is, there should be competition between objects from the current fixation and the objects from the previous fixation that are now in VWM and whose attentional weights are encapsulated to ensure their consolidation into passive VWM. This encapsulation should, however, only take place for objects that have been encoded into VWM and that are task-relevant (see the above description of TRAM).

The aim of **Study 3** (Poth & Schneider, 2016b, submitted) was to distinguish between the three hy-

potheses and thus to investigate whether objects compete for recognition across saccades. For this purpose, we performed two experiments. In both experiments, participants made saccades to an object (a red circle) in the visual periphery. Their task was to report a letter that was presented after the saccade within this object and terminated by a mask. The letter was shown for different durations. This allowed us to assess the processing speed of this letter by modeling report performance as a function of letter presentation duration (for a recent overview, see Bundesen et al., 2015). Before the saccade, either no, two, or four additional non-target objects (digits) were shown until the saccade was initiated. The non-target objects were in a similar color than the saccade target object.

In Experiment 1, the non-target objects were task-irrelevant. Under such conditions, the speed with which the postsaccadic letter was processed did not differ depending on the number of presaccadic non-target objects. This result argues against the second hypothesis, which predicted that objects per se compete across the saccade for object recognition.

In Experiment 2, the presaccadic non-target objects were task-relevant because they had to be matched against a probe at the end of a trial. Here, the processing speed of the postsaccadic letter now decreased with increasing number of presaccadic non-targets. Cross-experiment analyses confirmed that the effect of the number of presaccadic non-targets on the processing speed of the postsaccadic letter was indeed greater in Experiment 2 than in Experiment 1. These findings provide evidence against the first hypothesis, stating that objects cannot compete for recognition across saccades. Instead, they lend support to the third hypothesis that was based on TRAM theory (Schneider, 2013), holding that objects compete for recognition across saccades, but only if they are task-relevant.

In TRAM (Schneider, 2013), the findings of Study 3 can be explained as follows. Before the saccade, the saccade target object was the most relevant object in the visual field, with the highest attentional weight, and should thus have entered VWM (cf. Deubel & Schneider, 1996; Wischnewski et al., 2009, 2010). However, as VWM allows to retain more than one object, some of the presaccadic non-target objects should have been encoded into VWM as well. In particular, the presaccadic non-target objects were of a similar color than the saccade target and they appeared with a sharp onset, which implies a high bottom-up salience (cf. Yantis, 1993). According to TVA (Bundesen, 1990), their attentional weights should thus have been relatively high, supporting their encoding into VWM. After the saccade, the saccade target object remained as it was, only that the letter appeared in it. For this object, object correspondence should have been established across the saccade (as in Study 1 and 2, Poth et al., 2015; Poth & Schneider, 2016a). However, this should not have been the case for the presaccadic non-target objects. These objects were extinguished across the saccade, which should have drastically changed their attentional weights. Therefore, for the representations of these objects in VWM, object correspondence across the saccade should have been broken. This should have lead to the encapsulation of attentional weights, but only if the objects in VWM connected to the attentional weights had been task-relevant. The reason for this is that encapsulation of attentional weights happens only for objects in VWM that are task-relevant, in order to ensure only their consolidation into passive VWM. As a consequence of the encapsulation of attentional weights, there should have been fewer resources available for processing the postsaccadic letter. This explains that the letter was processed more slowly.

As an alternative to encapsulation, one might suppose that the lower postsaccadic processing speed was solely due to the higher attentional weights of task-relevant presaccadic non-target objects compared to irrelevant ones. Arguing against this, however, one may assume that even though irrelevant non-target objects should have had lower attentional weights than task-relevant ones, their attentional weights should not have been close to

zero. As mentioned above, they appeared with a sharp onset and shared features with the saccade target, implying relatively high attentional weights (Bundesen, 1990). Therefore, if attentional competition took place across the saccade irrespective of encapsulation, one should expect an effect of the number of presaccadic non-target objects on post-saccadic processing speed also for irrelevant ones. That there was no such effect (in fact, it was more likely that the effect was absent, see Poth & Schneider, 2016b, submitted), might thus argue against this hypothesis.

In conclusion, the findings of Study 3 (Poth & Schneider, 2016b, submitted) reveal that objects compete for object recognition across saccadic eye movements, but only if the objects are task-relevant. This argues that the current task decides how processing resources for object recognition are allocated to different objects in successive visual processing episodes. Attentional competition between objects across processing episodes seems limited to task-relevant objects. This may reflect an intelligent compromise: The risk of impaired object recognition in the current processing episode is only taken for those objects whose continued processing serves the current task.

Chapter 4

Short-term recognition across visual processing episodes

4.1 Short-term recognition requires encoding into visual working memory in previous processing episodes

The previous chapter focused on the mechanisms that link the current visual processing episode to the immediately preceding one, in order to support object recognition. In contrast to object recognition, short-term recognition requires to establish links from the current to several processing episodes in the recent past, not only to the immediately preceding one. More specifically, short-term recognition across visual processing episodes bears the matching problem (see section 1.5). It requires to match an object from the current episode to object representations that have been acquired in several recent episodes. It is an open question whether short-term recognition is based on mechanisms that are also involved in object recognition. For instance, one may ask whether object recognition in one episode has to be accomplished for an object in order for it to be available for short-term recognition in later episodes.

In **Study 4** (Poth & Schneider, 2016c), we aimed to make a first approach to this issue. As argued above, VWM is one of the most important mechanisms for object recognition. Particularly, objects are assumed to be recognized and become available for report and other actions with their encoding into VWM (Bundesen, 1990; Bundesen et al., 2005; Schneider, 2013). In contrast to the function of object recognition, short-term recognition is not recognizing that an object belongs to a certain category but that it has been viewed recently (in a previous processing episode, for which we also called it "episodic short-term recognition", Poth & Schneider, 2016c). Here, we asked whether short-term recognition of an object requires that it has entered VWM at the time it was encountered. Alternatively, it could be sufficient to activate visual features and categories in visual long-term memory, which should happen for all objects in the visual field as an initial processing step (Bundesen, 1990; Cowan, 1988; Henderson & Anes, 1994; Oberauer, 2002).

For instance, in the first and unselective stage of processing in TVA (Bundesen, 1990), the objects in the visual field are assumed to activate representations of visual features and categories in long-term memory. Thereby, the sensory evidence that the objects have certain features or belong to certain categories is computed as the basis of further processing. Only afterwards, in the second and selective stage of processing, these categorizations can be encoded into VWM (i.e. this is TVA's race towards VWM, see section 2.1).

To investigate whether it is necessary to encode objects into VWM for their later short-term recognition, we developed a paradigm combining letter report and probe recognition (Poth & Schneider, 2016c). In Experiment 1, participants viewed displays of ten different letters in circular arrangement. They memorized them over a retention interval and then reported as many of them as they could. Ten letters were displayed because this number exceeds estimates of VWM capacity

(e.g., Dyrholm, Kyllingsbæk, Espeseth, & Bundesen, 2011; Luck & Vogel, 1997; Poth et al., 2014; Shibuya & Bundesen, 1988) and ensures that participants could not report all letters. After the letter report, a single probe letter was shown. Short-term recognition was assessed as participants' performance in indicating whether or not the probe matched one of the preceding ten letters. To assess whether the probed letter had been encoded into VWM, the identity of the probe was conditionalized on the letters participants had reported. There were three conditions. The probe either matched one of the ten letters and had also been reported, it matched one of the ten letters but had not been reported, or it did not match any of the ten letters. In general, objects are assumed to be available for report and the control of other actions only if they have entered VWM (Bundesen, 1990; Bundesen et al., 2005; Martens & Wyble, 2010; Schneider, 2013). Therefore, we assumed that the letters participants reported had been encoded into VWM, whereas the letters they did not report had not reached VWM.

It is important to note that short-term recognition here referred to letters not from the immediately preceding visual processing episode, but from several processing episodes ago. The reason for this is that the display of the ten letters was followed by a retention interval (without letters) and by the action of reporting letters. Both of these phases of the task should have triggered a new visual processing episode due to changes in processing demands (Schneider, 2013).

The results of Experiment 1 of Study 4 (Poth & Schneider, 2016c) showed that participants' probe recognition performance was higher for reported letters than for those that had not been reported or not been shown. This is consistent with the hypothesis that short-term recognition presupposes encoding into VWM. However, reporting the letters could itself have effects on a later short-term recognition irrespective of encoding into VWM. For instance, reporting letters itself could interfere with visual features and categories in long-term memory

that had been activated by other letters. Without an intervening report, short-term recognition based on these initially activated features and categories might still be possible.

Additional analyses of letter reports in Experiment 1 revealed that participants predominantly reported letters that had been close to each other within the displays of the ten letters. This suggests that letters were encoded into VWM in a spatially-clustered manner. Experiment 2 made use of this finding to address the alternative explanation of the higher probe recognition performance for reported letters in Experiment 1.

The paradigm of Experiment 2 was similar to the one of Experiment 1. In contrast to Experiment 1, however, the ten letters were placed in colored frames, whereby the frame of one letter differed from the other frames. Participants only reported this highlighted letter. The probe letter either referred to this letter, to one of the two letters next to it, or two one of two letters on the other side of the letter display. Based on the spatially clustered encoding into VWM in Experiment 1, we assumed that letters next to the highlighted letter should have a higher probability of entering VWM than those on the other side of the display. Therefore, if short-term recognition required encoding into VWM, probe recognition performance should be higher for the letters near compared with those far away from the highlighted letter. The results supported this prediction. Probe recognition was higher for letters that had been near to compared with far away from the highlighted letter. Importantly, probe recognition performance was at chance level for these far away letters. This indicates that short-term recognition could be performed for letters that had been probable to enter VWM, but was impossible for letters with a lower probability of reaching VWM.

In sum, these findings argue that short-term recognition is restricted to those objects that have reached VWM. As such, the activation of features and categories in long-term memory seems insufficient to enable this function. This means that

VWM contributes to the solution of the matching problem of short-term recognition across visual processing episodes. That is, VWM seems to restrict the amount of information that has to be considered by the mechanisms performing the actual matching, the comparison of objects from the current and from recently past episodes. Furthermore, the findings indicate that VWM not only underlies the function of object recognition but also plays a crucial role for short-term recognition. This also implies that object recognition and short-term recognition share a common limit: Information about only a small number of objects can be encoded into capacity-limited VWM per visual processing episode to accomplish the two functions of visual cognition.

4.2 Priority in visual working memory impacts on distinct components of short-term recognition

As we have just seen, VWM seems to play an important role in both, object recognition and short-term recognition. As described before (see section 2.1), encoding into VWM is mediated by mechanisms of visual attention, selecting among all available object those that are important to the current task, thus prioritizing the objects for recognition (e.g., Bundesen, 1990; Duncan, 2006; Schneider, 1995; for an overview, see Poth & Schneider, 2013). There has been extensive research on such a prioritization of object information up to the time of VWM encoding (as reviewed by Bundesen et al., 2015; Duncan, 2006). However, a flexible use and application of visual information in task-driven behavior, requires that processing priorities can also be changed after VWM encoding. This may be especially important for accommodating and preparing for the requirements of upcoming visual processing episodes, such as an impending short-term recognition.

Prioritization of objects within VWM can be studied by means of the so-called *retro-cuing paradigm* (Griffin & Nobre, 2003, see also Landman, Spekreijse, & Lamme, 2003). In this experimental paradigm, participants memorize a set of visual objects over a retention interval, after which a probe object is shown. Participants then indicate whether the probe matches one of the objects that had been shown before. Short-term recognition is then assessed as performance in the probe recognition task (cf. Study 4, Poth & Schneider, 2016c). Importantly, a *retro-cue* (i.e. a "retrodictive cue") is presented after the display of objects but before the probe. In the variants of the paradigm that are of current interest, the retro-cue is either valid or neutral (e.g., Astle, Summerfield, Griffin, & Nobre, 2012; Kuo, Stokes, & Nobre, 2012). Valid retro-cues predict which of the memorized objects is going to be relevant for the upcoming comparison with the probe. Neutral retro-cues do not provide any information about the comparison. The main finding of this paradigm is that probe recognition performance is improved by valid as compared to neutral retro-cues. Short-term recognition thus benefits from the valid retro-cues.

Fueling a constant debate, retro-cues may impact on a number of different mechanisms to support performance in an upcoming short-term recognition task (for a review, see Souza & Oberauer, 2016). Most accounts assume that retro-cues affect memory-retention, by interacting with representations of the memorized objects in VWM (e.g., Kuo, Yeh, Chen, & D'Esposito, 2011; Lepsien, Thornton, & Nobre, 2011; Matsukura, Luck, & Vecera, 2007; Murray, Nobre, Clark, Cravo, & Stokes, 2013; Nobre, Griffin, & Rao, 2008, for a more extensive discussion, see Poth & Schneider, 2016d, submitted). In contrast to this view, one may, however, also hypothesize that retro-cues exert their effects by interacting with processing of the probe, improving its utility for short-term recognition. Moreover, the two hypotheses are not mutually exclusive, retro-cues could well influence memory-retention and processing of the probe.

In **Study 5** (Poth & Schneider, 2016d, submitted), we investigated these hypotheses in an experiment modifying the retro-cuing paradigm (and in a near-exact replication experiment). Participants

viewed two objects (colored squares) and memorized them over a retention interval. Afterwards, a probe was shown that either did or did not match one of the preceding objects. Short-term recognition was assessed as participants' performance in indicating if the probe did or did not match a preceding object. A retro-cue was presented in between the retention interval and the probe. The retro-cue was either valid or neutral. A valid retro-cue pointed at the location of the one of the preceding objects that was going to be relevant for the upcoming comparison with the probe. Specifically, if the probe matched a preceding object, it was always the one highlighted by the retro-cue. A neutral retro-cue did not contain such location information. Critically, the presentation duration of the probe was varied and it was terminated by a mask. This allowed us to assess performance in the short-term recognition task as a function of the presentation duration of the probe. We fit this data with an exponential model (Bundesen, 1990; Wickelgren, 1977) that disentangled two components of performance (among an additional component that is not relevant here, see Poth & Schneider, 2016d, submitted). The first component is the level of asymptotic performance which is reached when the probe is shown for a relatively long duration. Based on a number of findings, we can assume that perceptual processing of the probe improves with increasing presentation duration (e.g., Bundesen & Harms, 1999; Petersen & Andersen, 2012; Shibuya & Bundesen, 1988). At the asymptote, however, performance stops to increase with increasing presentation duration of the probe. At this point, perceptual processing of the probe should be over. Therefore, variations of the asymptote should not reflect processing of the probe but measure the performance in retaining the objects in VWM. The second component is the rate at which performance increases with increasing presentation duration of the probe (toward asymptotic performance). This is a well-established measure of processing speed (Bundesen, 1990; Wickelgren, 1977; see also Study 3, Poth & Schneider, 2016b, submitted). Here, it should represent the speed with

which the probe is processed for comparison with the objects in VWM in order to accomplish short-term recognition of the probe.

The results of Study 5 (Poth & Schneider, 2016d, submitted) showed that both of the two components were affected by the retro-cues. Valid retro-cues improved memory-retention, as assessed by the asymptotic level of performance. This is well in line with several accounts from the literature, assuming that retro-cues modulate object representations in VWM (Kuo et al., 2011; Lepsien et al., 2011; Matsukura et al., 2007; Murray et al., 2013; Nobre et al., 2008). However, valid retro-cues also accelerated processing of the probe, which is a finding that calls for an extension of current accounts. One interpretation of this effect is that briefer presentations of the probe caused it to be represented with low quality. The valid retro-cue may have compensated for this low quality. For instance the valid retro-cue could have strengthened of the cued object in VWM (Kuo et al., 2011; Lepsien et al., 2011; Nobre et al., 2008), so that the comparison could still be conducted. Shorter presentation durations of the probe (after exceeding a minimum presentation duration, see Poth & Schneider, 2016d, submitted) would still be sufficient for short-term recognition. Thus, functionally, this would be equivalent to an increase in the processing speed of the probe in the context of this task.

Alternatively, valid retro-cues may have accelerated processing of the probe by engaging a monitoring mechanism (see Poth & Schneider, 2016d, submitted, for further discussion). That is, once the valid retro-cue indicated which of the objects in VWM was going to be relevant for the upcoming short-term recognition task, the display could have been monitored for the appearance of the features of the cued object. This could have increased visual processing speed by means of the pigeonholing mechanism of TVA (Bundesen, 1990; see section 2.1). The perceptual decision bias for categorizing the upcoming probe as having the features of the cued object could have been set high. This

would have increased processing of these categorizations. However, such a monitoring by pigeonholing only works on trials on which the probe indeed matched the cued object. This problem may be solved by a setting a deadline for processing the probe. If the probe was not categorized as having the features of the cued object until this deadline, it would be decided that the probe did not match the cued object. Since valid retro-cues speeded up processing of the probe, the deadline could be set shorter compared with neutral retro-cues.

Together, the findings of Study 5 (Poth & Schneider, 2016d, submitted) demonstrate that prioritizing information in VWM can prepare the use of this information in an upcoming short-term recognition task. This seems to involve both, enhanced memory-retention in VWM and accelerated processing of the probe on which short-term recognition has to be performed. This suggests that priorities in VWM contribute to the solution of the matching problem of short-term recognition. They influence the retention of information from recently past episodes and the acquisition of new information in upcoming episodes. In this way, priorities in VWM may determine which and how much information from different processing episodes is taken into account by the mechanisms performing the actual matching of objects for short-term recognition.

The findings may add an interesting avenue to short-term recognition across visual processing episodes. They may suggest that task-driven information processing continues after information has entered VWM and that it takes new information (such as from a retro-cue) into account. Importantly, this may then prepare processing in the next visual processing episode, at least when this episode belongs to the same short-term recognition task.

Chapter 5

Episodic visual cognition: Discussion and outlook

The central question of this dissertation was how the two functions of visual cognition, object recognition and short-term recognition, are fulfilled across visual processing episodes. More specifically, the dissertation asked how the mechanisms underlying the two functions solve particular problems arising from the dissection of visual processing into episodes.

5.1 Mechanisms of episodic visual cognition for object and short-term recognition

It was argued that visual processing episodes pose two contrasting problems for the two functions of visual cognition. For object recognition, visual processing episodes lead to the problem of selective integration. That is, it must be decided whether object representations from two successive episodes should be integrated into a common or separated into distinct representations. Integration enables a cumulative processing of external objects across visual processing episodes (see e.g., Demeyer et al., 2009; Kahneman et al., 1992; Rayner et al., 1980). This is necessary for object recognition in situations where processing episodes are shorter than the processing time required for object recognition (see section 1.5). In addition, integration may support object recognition by preventing a competition for object recognition that would arise between two separate representations (Schneider, 2013). Separation is necessary to notice and discriminate changes of objects across episodes (cf. Deubel et al., 1996, 2002; Weiß et al., 2015). Such change perception is also

required in many situations of visually-guided behavior (e.g., Rensink, 2002).

In contrast to object recognition, for short-term recognition visual processing episodes lead to a problem of matching. While the selective integration problem refers to object recognition across one and the next episode, the matching problem of short-term recognition requires links between the current and (potentially) multiple episodes of the recent past (and irrespective of object recognition in the sense of categorization).

In our studies, we investigated components of the mechanisms underlying object and short-term recognition that may play key roles in solving the problems of visual processing episodes.

Our first three studies focused on object recognition across processing episodes of eye fixations separated by saccadic eye movements, because fixations are a type of episode ubiquitous in human vision (e.g., Gegenfurtner, 2016; Rolfs, 2015; Schütz et al., 2011).

In Study 1 (Poth et al., 2015) and Study 2 (Poth & Schneider, 2016a), we investigated a mechanism testing for object correspondence ("object continuity") across visual processing episodes, which has been put forward in TRAM theory (Schneider, 2013). Our findings indicated that such a test for correspondence between objects across saccades impacts on the perception of transsaccadic object displacements as well as on postsaccadic object recognition. That is, breaking object correspondence across the saccade improved

discrimination of transsaccadic object displacements but impaired postsaccadic object recognition. Moreover, we found that object recognition was impaired when object correspondence was broken by changing spatiotemporal object features across the saccade (by briefly blanking the object), but also by changing multiple surface features (contrast-polarity and associated luminance, color-and-luminance, and color alone). Based on TRAM (Schneider, 2013), we interpreted the findings as evidence that breaking object correspondence across the saccade causes presaccadic and postsaccadic representations of an external object to be kept as separate entities. This should allow comparisons between them, explaining the improved displacement perception. However, the separate representations should also compete for visual processing resources necessary for object recognition, explaining why object recognition was impaired. Conversely, establishing object correspondence across the saccade should lead to the integration of the presaccadic and postsaccadic object representations. Besides (or in addition to) preventing a competition for visual processing resources, this could improve object recognition by enabling the cumulative processing of the object across the saccade. Thus, taken together, the object correspondence mechanism proposed by TRAM theory may solve the selective integration problem of object recognition across saccades and thereby impact on object recognition in several ways.

An object correspondence mechanism may link successive episodes for visual processing of a single external object. For different external objects, however, visual processing episodes create another problem. Namely, limited visual processing resources (e.g., neurons; Bundesen et al., 2005; Desimone & Duncan, 1995) that are required for object recognition must be allocated intelligently to different objects in successive processing episodes.

In Study 3 (Poth & Schneider, 2016b, submitted), we examined how visual processing resources are distributed across different objects of successive episodes of eye fixations. Here, we found that objects compete for visual processing resources across the saccade separating the fixations, but only if they are task-relevant. Currently relevant objects of one fixation seem to cut the resources for object recognition in the next fixation, which becomes manifest in slower visual processing. These findings offer support for a key prediction of TRAM theory (Schneider, 2013), namely that the task-relevance of objects determines the distribution of visual processing resources not only within a fixation (as in classic attention theories Bundesen, 1990; Duncan & Humphreys, 1989; Wolfe, 1994) but also across intervening saccadic eye movements. In sum, this suggests that the distribution of processing resources for object recognition within and across processing episodes is likewise controlled by the current task. In this fashion, processing of different objects within and across episodes is reconciled in accordance with the task, which in turn implies that the task (or task-step) unites successive episodes under one roof of common processing goals (cf. Duncan, 2013).

Within a visual processing episode, processing for object recognition is assumed to end with the encoding of objects into VWM (e.g., Bundesen, 1990; Bundesen et al., 2005). Once encoded into VWM, objects are available for being reported (e.g., Bundesen, 1990; Luck & Vogel, 2013; Eriksson et al., 2015) or used for controlling goal-directed action (e.g., Schneider, 2013).

In Study 4 (Poth & Schneider, 2016c), we asked whether such an encoding of object information into VWM is also required beyond the current episode. More specifically, we asked if object information must have been processed up to the level of VWM in the episode it was acquired, in order to be available for short-term recognition in later episodes. Our results seem to support this notion. We found that objects that supposedly did not reach VWM (i.e. had not been reported or were unlikely to be reported) were not available for performing a subsequent short-term recognition task. Therefore, encoding object information into VWM seems to be a processing step that prepares

and is necessary for short-term recognition in later processing episodes. In this manner, VWM contributes to the solution of the matching problem of short-term recognition across visual processing episodes. That is, VWM seems to limit the amount of information that has to be considered by the mechanisms performing the actual matching, the comparison of objects from the current and from recently past episodes.

Study 5 (Poth & Schneider, 2016d, submitted) continued the investigation of VWM as a mechanism underlying short-term recognition across processing episodes. Here, we examined how prioritizing objects within VWM can prepare for short-term recognition in an upcoming processing episode. In particular, we used a paradigm that allowed to disentangle two components of short-term recognition performance. The first component is memory-retention in VWM. The second one is the processing speed of a probe object that had to be matched against objects held in VWM, in order to perform the short-term recognition. Prioritization in VWM was studied by retro-cues, which did or did not indicate the object in VWM that was going to be relevant for short-term recognition. We found both of the two components affected by the prioritization. Prioritizing an object in VWM improved its memory-retention, but also accelerated the processing of a probe object in the upcoming short-term recognition task. This suggests that the matching problem of short-term recognition across visual processing episodes is solved in a dynamic and task-driven fashion. That is, current processing priorities contribute to the problem solution in two ways. First, by influencing the retention of information of recently past episodes, and second, by guiding the acquisition of new information in upcoming episodes. In this fashion, the priorities determine which and how much information from different processing episodes is used by the mechanisms performing the actual matching for short-term recognition.

Taken together, the findings of the discussed studies reveal specific mechanisms that contribute to object recognition and short-term recognition across visual processing episodes. The mechanism of object correspondence seems to link one processing episode to the next for encoding into VWM and thus for object recognition. The mechanism of VWM seems implied in both, object recognition in the current episode, and in the preparation of short-term recognition in later episodes. Mechanisms that set priorities within VWM seem to impact on the next processing episodes by influencing memory-retention as well as the acquisition of new information in service of short-term recognition.

However, by introducing these mechanisms in the context of visual processing episodes, the studies also raise a number of new questions for future research.

5.2 Open questions of episodic visual cognition

As explained in the Introduction (see section 1.4), visual processing episodes arise due to a number of different factors (Schneider, 2013). We here examined the mechanisms of object and short-term recognition across different types of visual processing episodes.

Our first three studies (Poth et al., 2015; Poth & Schneider, 2016a, 2016b, submitted) focused on eye fixations as episodes that must be overcome for object recognition. Fixations are separated by saccadic eye movements, which might make them special in the sense that information about their onset-time and direction is available in advance. For instance, the saccade-eliciting signals can be made available to brain centers for visual processing by corollary discharges (see e.g., Wurtz, 2008; Wurtz et al., 2011), and this information may contribute to the linking of successive fixations (e.g., by specifying a time-window in which the test for transsaccadic object correspondence, Poth et al., 2015; Poth & Schneider, 2016a, should be conducted). Such predictive information would be missing for visual processing episodes triggered by the appearance or disappearance of objects in the visual field (for a discussion, see Schneider, 2013), or by the occlusion or movement of objects (Hollingworth &

Franconeri, 2009; Kahneman et al., 1992).

In addition, a different kind of predictive information about upcoming processing episodes may be available when the episode is embedded in a known task, such as in the presented studies (cf. Duncan, 2013). For these reasons, an interesting avenue for future research may be to investigate how the factors eliciting and predicting visual processing episodes impact on the solutions put forward by mechanisms of visual cognition to overcome them.

5.3 Is episodicness a principle of visual cognition?

Up to this point, we have considered visual processing episodes as a problem that must be overcome for visual cognition, specifically for its functions of object and short-term recognition. However, the final part of this dissertation synopsis aims to take a different view and speculate about the functional value that processing episodes might have for visual cognition.

The object correspondence mechanism proposed by TRAM theory (Schneider, 2013) presents a solution to the selective integration problem of object recognition across visual processing episodes. At the beginning of each visual processing episode, the mechanism tests whether current external objects correspond to objects from the previous episode (that are now in VWM). Depending on the outcome of this test, information about an external object is either integrated or separated across the visual processing episodes.

Integrating information about an object over time into one representation may be a strategy for fast information processing, because it can proceed cumulatively (e.g., Demeyer et al., 2009; Kahneman et al., 1992). It should also prevent attentional competition for object recognition, which arose if separate representations were formed (Schneider, 2013). Integration may also provide the most robust object representations (e.g., by capitalizing on potential information redundancy, as supposed for information from different senses, Ernst

& Bülthoff, 2004). It may also allow a sparse and computationally efficient processing (cf. Olshausen & Field, 2004), because only one instead of two (or more) representations has to be dealt with. However, if information about two actually different states of an object or two different objects is integrated, then the differences would be missed. This may be the case because there are no two representations that can be compared to detect the difference (as discussed for the perception of visual stability, see section 3.1).

Whether it is more important to have robust and sparse object representations or to be able to detect and discriminate object changes over time depends on the current task and environmental circumstances. Critically, whether the one or the other requirement prevails may change quickly. This thought may lead to the new view that having visual processing episodes can also support visual cognition and not only challenge it. That is, dissecting visual processing into distinct episodes offers temporally regular and controllable checkpoints at which it can be decided whether object information should be integrated or separated. Therefore, having episodes rather than a continuous stream of visual processing may reflect a compromise between the representational robustness provided by integration and the temporal resolution for detecting object changes provided by separation.

In addition to such a compromise between robustness and temporal resolution of representations, visual processing episodes may enhance processing by limiting processing demands. Specifically, they may restrict the amount of information that is carried forward in time. For instance, Study 4 (Poth & Schneider, 2016c) suggested that initial processing steps that activate for all object in the visual field specific visual features and categories in long-term memory are not sufficient for an object's later short-term recognition. Instead, processing the object up to the level of VWM seems to be required. This might hint at that the lingering activation of visual features and categories is

not sustained over time (and hence cannot be used for short-term recognition). As a consequence, there may be less interference between features and lower processing demands at the initial steps of processing (e.g., the formation of proto-objects that is proposed by TRAM theory, Schneider, 2013, see also Wischnewski et al., 2009, 2010). Dissecting processing into episodes and clearing initial stages of processing at their beginning may be more efficient than resolving conflicts of current and past feature activations. This should support object as well as short-term recognition. In addition, this proposal is also in line with the older idea that carrying only a limited amount of information forward in time is beneficial for action control (as in the selection-for-action view of attention, Allport, 1987; see also Neumann, 1987). That is, the parameters for controlling action may be obtained faster (cf. Neumann, 1987), because they can be extracted from pre-restricted (e.g., in VWM) information instead of all available information (e.g., at the retina).

To conclude, visual processing episodes can be regarded from two points of view. On the one hand, there are a number of external factors and task-requirements that result in visual processing episodes (Schneider, 2013) and that must be addressed by mechanisms of visual cognition. Otherwise functions such as object and short-term recognition were not realizable. On the other hand, however, "episodicness" may also constitute a principle of processing that finds expression in different mechanisms, and reconciles representational robustness, temporal resolution, and potential interference of information.

References

Albright, T. D. (2013). High-level visual processing: Cognitive influences. In E. R. Kandel, J. H. Schwartz, T. L. Jessel, S. A. Siegelbaum, & A. J. Hudspeth (Eds.), *Principles of Neural Science* (pp. 621–637). New York: McGraw-Hill.

Allport, D. A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on Perception and Action* (pp. 395–419). London: Lawrence Erlbaum Associates.

Astle, D. E., Summerfield, J., Griffin, I., & Nobre, A. C. (2012). Orienting attention to locations in mental representations. *Attention, Perception & Psychophysics*, *74*(1), 146–162. doi: 10.3758/s13414-011-0218-3

Bays, P. M. (2015). Spikes not slots: Noise in neural populations limits working memory. *Trends in Cognitive Sciences*, *19*(8), 431–438. doi: 10.1016/j.tics.2015.06.004

Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience*, *33*, 1–21. doi: 10.1146/annurev-neuro-060909-152823

Block, N. (1995). How many concepts of consciousness? *Behavioral and Brain Sciences*, *18*(02), 272. doi: 10.1017/S0140525X00038486

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences*, *15*(12), 567–575. doi: 10.1016/j.tics.2011.11.001

Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research*, *15*(6), 719–722. doi: 10.1016/0042-6989(75)90290-4

Bridgeman, B., van der Heijden, A. H. C., & Velichkovsky, B. M. (1994). A theory of visual stability across saccadic eye movements. *Behavioral and Brain Sciences*, *17*(02), 247–292. doi: 10.1017/S0140525X00034361

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523–547. doi: 10.1037/0033-295X.97.4.523

Bundesen, C., & Habekost, T. (2008). *Principles of Visual Attention: Linking Mind and Brain*. Oxford, UK: Oxford University Press.

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review*, *112*(2), 291–328. doi: 10.1037/0033-295X.112.2.291

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2011). A neural theory of visual attention and short-term memory (NTVA). *Neuropsychologia*, *49*(6), 1446–1457. doi: 10.1016/j.neuropsychologia.2010.12.006

Bundesen, C., & Harms, L. (1999). Single-letter recognition as a function of exposure duration. *Psychological Research*, *62*(4), 275–279. doi: 10.1007/s004260050057

Bundesen, C., Vangkilde, S., & Petersen, A. (2015). Recent developments in a computational theory of visual attention (TVA). *Vision Research*, *116*, 210–218. doi: 10.1016/j.visres.2014.11.005

Cavanagh, P. (2011). Visual cognition. *Vision Research*, *51*(13), 1538–1551. doi: 10.1016/j.visres.2011 .01.015

Cavanagh, P., Hunt, A. R., Afraz, A., & Rolfs, M. (2010). Visual stability based on remapping of attention pointers. *Trends in Cognitive Sciences*, *14*(4), 147–153. doi: 10.1016/j.tics.2010.01.007

Cavanaugh, J., Berman, R. A., Joiner, W. M., & Wurtz, R. H. (2016). Saccadic Corollary Discharge Underlies Stable Visual Perception. *Journal of Neuroscience*, *36*(1), 31–42. doi: 10.1523/ JNEUROSCI.2054-15.2016

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*(2), 163–191. doi: 10.1037/0033-2909.104.2.163

Curcio, C. A., & Allen, K. A. (1990). Topography of ganglion cells in human retina. *The Journal of Comparative Neurology*, *300*(1), 5–25. doi: 10.1002/cne.903000103

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204–211. doi: 10.1016/j.tics.2006.03.007

Demeyer, M., de Graef, P., Wagemans, J., & Verfaillie, K. (2009). Transsaccadic identification of highly similar artificial shapes. *Journal of Vision*, *9*(4), 28.1-14. doi: 10.1167/9.4.28

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222. doi: 10.1146/annurev.ne.18.030195.001205

Deubel, H., Bridgeman, B., & Schneider, W. X. (1998). Immediate post-saccadic information mediates space constancy. *Vision Research*, *38*(20), 3147–3159. doi: 10.1016/S0042-6989(98)00048-0

Deubel, H., & Schneider, W. X. (1994). Perceptual stability and postsaccadic visual information: Can man bridge a gap? *Behavioral and Brain Sciences*, *17*(02), 259. doi: 10.1017/ S0140525X00034397

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*(12), 1827–1837. doi: 10.1016/0042 -6989(95)00294-4

Deubel, H., Schneider, W. X., & Bridgeman, B. (1996). Postsaccadic target blanking prevents saccadic suppression of image displacement. *Vision Research*, *36*(7), 985–996. doi: 10.1016/ 0042-6989(95)00203-0

Deubel, H., Schneider, W. X., & Bridgeman, B. (2002). Transsaccadic memory of position and form. In J. Hyönä, D. P. Munoz, W. Heide, & R. Radach (Eds.), (Vol. 140, pp. 165–180). Amsterdam, NL: Elsevier.

Duhamel, Colby, C., & Goldberg, M. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, *255*(5040), 90–92. doi: 10.1126/science .1553535

Duncan, J. (2006). EPS Mid-Career Award 2004: Brain mechanisms of attention. *Quarterly Journal of Experimental Psychology*, *59*(1), 2–27. doi: 10.1080/17470210500260674

Duncan, J. (2013). The structure of cognition: Attentional episodes in mind and brain. *Neuron*, *80*(1), 35–50. doi: 10.1016/j.neuron.2013.09.015

Duncan, J., Bundesen, C., Olson, A., Humphreys, G., Chavda, S., & Shibuya, H. (1999). Systematic analysis of deficits in visual attention. *Journal of Experimental Psychology: General*, *128*(4), 450–478. doi: 10.1037/0096-3445.128.4.450

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458. doi: 10.1037/0033-295X.96.3.433

Dyrholm, M., Kyllingsbæk, S., Espeseth, T., & Bundesen, C. (2011). Generalizing parametric models by introducing trial-by-trial parameter variability: The case of TVA. *Journal of Mathematical Psychology*, *55*(6), 416–429. doi: 10.1016/j.jmp.2011.08.005

Eriksson, J., Vogel, E. K., Lansner, A., Bergstrom, F., & Nyberg, L. (2015). Neurocognitive architecture of working memory. *Neuron*, *88*(1), 33–46. doi: 10.1016/j.neuron.2015.09.020

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*(4), 162–169. doi: 10.1016/j.tics.2004.02.002

Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: a priority map for target selection. *Trends in Cognitive Sciences*, *10*(8), 382–390. doi: 10.1016/j.tics.2006.06.011

Finke, K., Bublak, P., Dose, M., Müller, H. J., & Schneider, W. X. (2006). Parameter-based assessment of spatial and non-spatial attentional deficits in Huntington's disease. *Brain*, *129*(5), 1137–1151. doi: 10.1093/brain/awl040

Finke, K., Bublak, P., Krummenacher, J., Kyllingsbæk, S., Müller, H. J., & Schneider, W. X. (2005). Usability of a theory of visual attention (TVA) for parameter-based measurement of attention I: Evidence from normal subjects. *Journal of the International Neuropsychological Society*, *11*(07). doi: 10.1017/S1355617705050976

Flombaum, J. I., Scholl, B. J., & Santos, L. R. (2009). Spatiotemporal priority as a fundamental principle of object persistence. In B. Hood & L. R. Santos (Eds.), *The Origins of Object Knowledge* (pp. 135–164). Oxford, UK: Oxford University Press.

Foerster, R. M., Carbone, E., Koesling, H., & Schneider, W. X. (2011). Saccadic eye movements in a high-speed bimanual stacking task: Changes of attentional control during learning and automatization. *Journal of Vision*, *11*(7), 9. doi: 10.1167/11.7.9

Foerster, R. M., Poth, C. H., Behler, C., Botsch, M., & Schneider, W. X. (2016). Using the virtual reality device Oculus Rift for neuropsychological assessment of visual processing capabilities. *Scientific Reports*, *6*, 37016. doi: 10.1038/srep37016

Gegenfurtner, K. R. (2003). Cortical mechanisms of colour vision. *Nature Reviews Neuroscience*, *4*(7), 563–572. doi: 10.1038/nrn1138

Gegenfurtner, K. R. (2016). The interaction between vision and eye movements. *Perception*. doi: 10.1177/0301006616657097

Gegenfurtner, K. R., & Kiper, D. C. (2003). Color vision. *Annual Review of Neuroscience*, *26*, 181–206. doi: 10.1146/annurev.neuro.26.041002.131116

Gilbert, C. D. (2013a). The constructive nature of visual processing. In E. R. Kandel, J. H. Schwartz, T. L. Jessel, S. A. Siegelbaum, & A. J. Hudspeth (Eds.), *Principles of Neural Science* (pp. 556–576). New York: McGraw-Hill.

Gilbert, C. D. (2013b). Intermediate-level visual processing and visual primitives. In E. R. Kandel, J. H. Schwartz, T. L. Jessel, S. A. Siegelbaum, & A. J. Hudspeth (Eds.), *Principles of Neural Science* (pp. 602–620). New York: McGraw-Hill.

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, *15*(8), 1176–1194. doi: 10.1162/089892903322598139

Habekost, T. (2015). Clinical TVA-based studies: a general review. *Frontiers in Psychology*, *6*, 290. doi: 10.3389/fpsyg.2015.00290

Habekost, T., Petersen, A., & Vangkilde, S. (2014). Testing attention: comparing the ANT with TVA-based assessment. *Behavior Research Methods*, *46*(1), 81–94. doi: 10.3758/s13428-013-0341-2

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, *9*(4), 188–194. doi: 10.1016/j.tics.2005.02.009

Henderson, J. M., & Anes, M. D. (1994). Roles of object-file review and type priming in visual identification within and across eye fixations. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(4), 826–839. doi: 10.1037/0096-1523.20.4.826

Herwig, A., & Schneider, W. X. (2014). Predicting object features across saccades: Evidence from object recognition and visual search. *Journal of Experimental Psychology: General*, *143*(5), 1903–1922. doi: 10.1037/a0036781

Hibino, H. (1992). Red-green and yellow-blue opponent-color responses as a function of retinal eccentricity. *Vision Research*, *32*(10), 1955–1964. doi: 10.1016/0042-6989(92)90055-N

Higgins, E., & Rayner, K. (2015). Transsaccadic processing: Stability, integration, and the potential role of remapping. *Attention, Perception & Psychophysics*, *77*(1), 3–27. doi: 10.3758/s13414-014-0751-y

Hollingworth, A., & Franconeri, S. L. (2009). Object correspondence across brief occlusion is established on the basis of both spatiotemporal and surface feature cues. *Cognition*, *113*(2), 150–166. doi: 10.1016/j.cognition.2009.08.004

Hollingworth, A., Richard, A. M., & Luck, S. J. (2008). Understanding the function of visual short-term memory: transsaccadic memory, object correspondence, and gaze correction. *Journal of Experimental Psychology: General*, *137*(1), 163–181. doi: 10.1037/0096-3445.137.1.163

Irwin, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 307–317. doi: 10.1037/0278-7393.18.2.307

Irwin, D. E., Brown, J. S., & Sun, J.-s. (1988). Visual masking and visual integration across saccadic eye movements. *Journal of Experimental Psychology: General*, *117*(3), 276–287. doi: 10.1037/0096-3445.117.3.276

Irwin, D. E., & Gordon, R. D. (1998). Eye movements, attention and trans-saccadic memory. *Visual Cognition*, *5*(1-2), 127–155. doi: 10.1080/713756783

Johnson, M. A. (1986). Color vision in the peripheral retina. *American Journal of Optometry & Physiological Optics*, *63*, 97–103.

Kahana, M. J., & Sekuler, R. (2002). Recognizing spatial patterns: A noisy exemplar approach. *Vision Research*, *42*(18), 2177–2192. doi: 10.1016/S0042-6989(02)00118-9

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219. doi: 10.1016/0010-0285(92)90007-O

Kastner, S., & Pinsk, M. A. (2004). Visual attention as a multilevel selection process. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(4), 483–500. doi: 10.3758/CABN.4.4.483

Krauzlis, R. J., Lovejoy, L. P., & Zenon, A. (2013). Superior colliculus and visual spatial attention. *Annual Review of Neuroscience*, *36*, 165–182. doi: 10.1146/annurev-neuro-062012-170249

Krock, R. M., & Moore, T. (2014). The influence of gaze control on visual perception: Eye movements and visual stability. *Cold Spring Harbor Symposia on Quantitative Biology*, *79*, 123–130. doi: 10.1101/sqb.2014.79.024836

Kuo, B.-C., Stokes, M. G., & Nobre, A. C. (2012). Attention modulates maintenance of representations in visual short-term memory. *Journal of Cognitive Neuroscience*, *24*(1), 51–60. doi: 10.1162/jocn_a_00087

Kuo, B.-C., Yeh, Y.-Y., Chen, A. J.-W., & D'Esposito, M. (2011). Functional connectivity during top-down modulation of visual short-term memory representations. *Neuropsychologia*, *49*(6), 1589–1596. doi: 10.1016/j.neuropsychologia.2010.12.043

Land, M. F., & Tatler, B. W. (2009). *Looking and acting: Vision and eye movements in natural behaviour*. Oxford, UK: Oxford University Press.

Landman, R., Spekreijse, H., & Lamme, V. A. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research*, *43*(2), 149–164. doi: 10.1016/S0042-6989(02)00402-9

Larocque, J. J., Lewis-Peacock, J. A., & Postle, B. R. (2014). Multiple neural states of representation in short-term memory? It's a matter of attention. *Frontiers in Human Neuroscience*, *8*, 5. doi: 10.3389/fnhum.2014.00005

Lepsien, J., Thornton, I., & Nobre, A. C. (2011). Modulation of working-memory maintenance by directed attention. *Neuropsychologia*, *49*(6), 1569–1577. doi: 10.1016/j.neuropsychologia.2011.03.011

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, *7*(11), 3416–3468.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. doi: 10.1038/36846

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391–400. doi: 10.1016/j.tics.2013.06.006

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356. doi: 10.1038/nn.3655

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York, USA: W. H. Freeman and Company.

Martens, S., & Wyble, B. (2010). The attentional blink: Past, present, and future of a blind spot in perceptual awareness. *Neuroscience and Biobehavioral Reviews*, *34*(6), 947–957. doi: 10.1016/j.neubiorev.2009.12.005

Mathôt, S., & Theeuwes, J. (2011). Visual attention and stability. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1564), 516–527. doi: 10.1098/rstb.2010.0187

Matsukura, M., Luck, S. J., & Vecera, S. P. (2007). Attention effects during visual short-term memory maintenance: Protection or prioritization? *Perception & Psychophysics*, *69*(8), 1422–1434. doi: 10.3758/BF03192957

Meister, M., & Tessier-Lavigne, M. (2013). Low-level visual processing: The retina. In E. R. Kandel, J. H. Schwartz, T. L. Jessel, S. A. Siegelbaum, & A. J. Hudspeth (Eds.), *Principles of Neural Science* (pp. 577–601). New York: McGraw-Hill.

Mitroff, S. R., & Alvarez, G. A. (2007). Space and time, not surface features, guide object persistence. *Psychonomic Bulletin & Review*, *14*(6), 1199–1204. doi: 10.3758/BF03193113

Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic theory of working memory. *Science*, *319*(5869), 1543–1546. doi: 10.1126/science.1150769

Moutoussis, K. (2015). The physiology and psychophysics of the color-form relationship: A review. *Frontiers in Psychology*, *6*, 1407. doi: 10.3389/fpsyg.2015.01407

Murray, A. M., Nobre, A. C., Clark, I. A., Cravo, A. M., & Stokes, M. G. (2013). Attention restores discrete items to visual short-term memory. *Psychological Science*, *24*(4), 550–556. doi: 10.1177/0956797612457782

Nagy, A. L., & Wolf, S. (1993). Red-green color discrimination in peripheral vision. *Vision Research*, *33*(2), 235–242. doi: 10.1016/0042-6989(93)90161-O

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on Perception and Action* (pp. 361–394). London: Lawrence Erlbaum Associates.

Neumann, O. (1990). Visual attention and action. In O. Neumann & W. Prinz (Eds.), *Relationships between Perception and Action* (pp. 227–267). Berlin: Springer.

Nobre, A. C., Griffin, I. C., & Rao, A. (2008). Spatial attention can bias search in visual short-term memory. *Frontiers in Human Neuroscience*, *1*, 4. doi: 10.3389/neuro.09.004.2007

Nordfang, M., Dyrholm, M., & Bundesen, C. (2013). Identifying bottom-up and top-down components of attentional weight by experimental analysis and computational modeling. *Journal of Experimental Psychology: General*, *142*(2), 510–535. doi: 10.1037/a0029631

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and Self-Regulation* (pp. 1–18). New York, NY, USA: Springer.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 411–421. doi: 10.1037//0278-7393.28.3.411

Olivers, C. N. L., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences*, *15*(7), 327–334. doi: 10.1016/j.tics.2011.05.004

Olshausen, B. A., & Field, D. J. (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, *14*(4), 481–487. doi: 10.1016/j.conb.2004.07.007

Oostwoud Wijdenes, L., Marshall, L., & Bays, P. M. (2015). Evidence for Optimal Integration of Visual Feature Representations across Saccades. *Journal of Neuroscience*, *35*(28), 10146–10153. doi: 10.1523/JNEUROSCI.1040-15.2015

Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. Lloyd (Eds.), *Cognition and Categorization* (pp. 259–303). Mahwah, NJ, USA: Lawrence Erlbaum Associates.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA, USA: MIT Press.

Petersen, A., & Andersen, T. S. (2012). The effect of exposure duration on visual character identification in single, whole, and partial report. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(2), 498–514. doi: 10.1037/a0026728

Petersen, A., Kyllingsbæk, S., & Bundesen, C. (2012). Measuring and modeling attentional dwell time. *Psychonomic Bulletin & Review*, *19*(6), 1029–1046. doi: 10.3758/s13423-012-0286-y

Petersen, A., Kyllingsbæk, S., & Bundesen, C. (2013). Attentional dwell times for targets and masks. *Journal of Vision*, *13*(3). doi: 10.1167/13.3.34.

Poth, C. H., Herwig, A., & Schneider, W. X. (2015). Breaking object correspondence across saccadic eye movements deteriorates object recognition. *Frontiers in Systems Neuroscience*, *9*, 176. doi: 10.3389/fnsys.2015.00176

Poth, C. H., Petersen, A., Bundesen, C., & Schneider, W. X. (2014). Effects of monitoring for visual events on distinct components of attention. *Frontiers in Psychology*, *5*, 930. doi: 10.3389/fpsyg.2014.00930

Poth, C. H., & Schneider, W. X. (2013). Aufmerksamkeit [Attention]. In A. Stephan & S. Walter (Eds.), *Handbuch Kognitionswissenschaft* (pp. 221–230). Weimar: J. B. Metzler.

Poth, C. H., & Schneider, W. X. (2016a). Breaking object correspondence across saccades impairs object recognition: The role of color and luminance. *Journal of Vision*, *16*(11), 1. doi: 10.1167/16.11.1

Poth, C. H., & Schneider, W. X. (2016b, submitted). Attentional competition across saccadic eye movements.

Poth, C. H., & Schneider, W. X. (2016c). Episodic short-term recognition requires encoding into visual working memory: Evidence from probe recognition after letter report. *Frontiers in Psychology*, *7*, 1440. doi: 10.3389/fpsyg.2016.01440

Poth, C. H., & Schneider, W. X. (2016d, submitted). Prioritization in visual working memory enhances memory-retention and speeds up processing in a comparison task.

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology*, *12*(2), 206–226. doi: 10.1016/0010-0285(80)90009-2

Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition*, *7*(1-3), 17–42. doi: 10.1080/135062800394667

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology*, *53*, 245–277. doi: 10.1146/annurev.psych.53.100901.135125

Rolfs, M. (2015). Attention in active vision: A perspective on perceptual continuity across saccades. *Perception*, *44*(8-9), 900–919. doi: 10.1177/0301006615594965

Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, *354*(6316), 1136–1139. doi: 10.1126/science.aah7011

Schall, J. D. (2009). Frontal eye fields. In M. D. Binder, N. Hirokawa, & U. Windhorst (Eds.), *Encyclopedia of Neuroscience* (pp. 1635–1638). Berlin, Germany: Springer.

Schneider, W. X. (1995). VAM: A neuro-cognitive model for visual attention control of segmentation, object recognition, and space-based motor action. *Visual Cognition*, *2*(2-3), 331–376. doi: 10.1080/13506289508401737

Schneider, W. X. (2013). Selective visual processing across competition episodes: a theory of task-driven visual attention and working memory. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(1628), 20130060. doi: 10.1098/rstb.2013.0060

Schneider, W. X., & Deubel, H. (2002). Selection-for-perception and selection-for-spatial-motor-action are coupled by visual attention: a review of recent findings and new evidence from stimulus-driven saccade control. In W. Prinz & B. Hommel (Eds.), *Common Mechanisms in Perception and Action: Attention and Performance XIX* (pp. 609–627). Oxford, UK: Oxford University Press.

Scholl, B. J. (2007). Object persistence in Philosophy and Psychology. *Mind & Language*, *22*(5), 563–591. doi: 10.1111/j.1468-0017.2007.00321.x

Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Eye movements and perception: A selective review. *Journal of Vision*, *11*(5). doi: 10.1167/11.5.9

Shibuya, H., & Bundesen, C. (1988). Visual selection from multielement displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(4), 591–600. doi: 10.1037/0096-1523.14.4.591

Sommer, M. A., & Wurtz, R. H. (2006). Influence of the thalamus on spatial visual processing in frontal cortex. *Nature*, *444*(7117), 374–377. doi: 10.1038/nature05279

Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception & Psychophysics*, *78*(7), 1839–1860. doi: 10.3758/s13414-016-1108-5

Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends in Cognitive Sciences*, *19*(7), 394–405. doi: 10.1016/j.tics.2015.05.004

Strasburger, H., Rentschler, I., & Juttner, M. (2011). Peripheral vision and pattern recognition: a review. *Journal of Vision*, *11*(5), 13. doi: 10.1167/11.5.13

Tas, A. C. (2015). *The role of visual stability in representations of pre- and post-saccadic objects* (PhD thesis, University of Iowa, Iowa, USA). Retrieved from `http://ir.uiowa.edu/etd/1916`.

Tas, A. C., Moore, C. M., & Hollingworth, A. (2012). An object-mediated updating account of insensitivity to transsaccadic change. *Journal of Vision*, *12*(11), 18. doi: 10.1167/12.11.18

Umeno, M. M., & Goldberg, M. E. (1997). Spatial processing in the monkey frontal eye field. I. Predictive visual responses. *Journal of Neurophysiology*, *78*, 1373–1383.

Walker, M. F., Fitzgibbon, E. J., & Goldberg, M. E. (1995). Neurons in the monkey superior colliculus predict the visual result of impending saccadic eye movements. *Journal of Neurophysiology*, *73*, 1988–2003.

Weiß, K., Schneider, W. X., & Herwig, A. (2015). A "blanking effect" for surface features: Transsaccadic spatial-frequency discrimination is improved by postsaccadic blanking. *Attention, Perception & Psychophysics*, *77*(5), 1500–1506. doi: 10.3758/s13414-015-0926-1

Wexler, M., & Collins, T. (2014). Orthogonal steps relieve saccadic suppression. *Journal of Vision*, *14*(2). doi: 10.1167/14.2.13

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, *41*(1), 67–85. doi: 10.1016/0001-6918(77)90012-9

Wischnewski, M., Belardinelli, A., Schneider, W. X., & Steil, J. J. (2010). Where to look next? Combining static and dynamic proto-objects in a TVA-based model of visual attention. *Cognitive Computation*, *2*(4), 326–343. doi: 10.1007/s12559-010-9080-1

Wischnewski, M., Steil, J. J., Kehrer, L., & Schneider, W. X. (2009). Integrating inhomogeneous processing and proto-object formation in a computational model of visual attention. In H. Ritter, G. Sagerer, R. Dillmann, & M. Buss (Eds.), *Human Centered Robot Systems* (Vol. 6, pp. 93–102). Berlin, Germany: Springer. doi: 10.1007/978-3-642-10403-9_10

Wittenberg, M., Bremmer, F., & Wachtler, T. (2008). Perceptual evidence for saccadic updating of color stimuli. *Journal of Vision*, *8*(14), 9.1-9. doi: 10.1167/8.14.9

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238. doi: 10.3758/BF03200774

Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Research*, *48*(20), 2070–2089. doi: 10.1016/j.visres.2008.03.021

Wurtz, R. H., Joiner, W. M., & Berman, R. A. (2011). Neuronal mechanisms for visual stability: Progress and problems. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1564), 492–503. doi: 10.1098/rstb.2010.0186

Yantis, S. (1993). Stimulus-driven attentional capture. *Current Directions in Psychological Science*, *2*(5), 156–161. doi: 10.1111/1467-8721.ep10768973

Zelinsky, G. J., & Bisley, J. W. (2015). The what, where, and why of priority maps and their interactions with visual working memory. *Annals of the New York Academy of Sciences*, *1339*, 154–164. doi: 10.1111/nyas.12606

Zhou, F., Kahana, M. J., & Sekuler, R. (2004). Short-term episodic memory for visual textures: A roving probe gathers some memory. *Psychological Science*, *15*(2), 112–118. doi: 10.1111/j.0963-7214.2004.01502007.x

Zirnsak, M., Steinmetz, N. A., Noudoost, B., Xu, K. Z., & Moore, T. (2014). Visual space is compressed in prefrontal cortex before eye movements. *Nature*, *507*(7493), 504–507. doi: 10.1038/nature13149

Original studies[1]

Study 1

Breaking object correspondence across saccadic eye movements deteriorates object recognition

Breaking Object Correspondence Across Saccadic Eye Movements Deteriorates Object Recognition

Christian H. Poth[1,2]*, Arvid Herwig[1,2] and Werner X. Schneider[1,2]

[1] Neuro-Cognitive Psychology, Department of Psychology, Bielefeld University, Bielefeld, Germany, [2] Cluster of Excellence Cognitive Interaction Technology, Bielefeld University, Bielefeld, Germany

Visual perception is based on information processing during periods of eye fixations that are interrupted by fast saccadic eye movements. The ability to sample and relate information on task-relevant objects across fixations implies that correspondence between presaccadic and postsaccadic objects is established. Postsaccadic object information usually updates and overwrites information on the corresponding presaccadic object. The presaccadic object representation is then lost. In contrast, the presaccadic object is conserved when object correspondence is broken. This helps transsaccadic memory but it may impose attentional costs on object recognition. Therefore, we investigated how breaking object correspondence across the saccade affects postsaccadic object recognition. In Experiment 1, object correspondence was broken by a brief postsaccadic blank screen. Observers made a saccade to a peripheral object which was displaced during the saccade. This object reappeared either immediately after the saccade or after the blank screen. Within the postsaccadic object, a letter was briefly presented (terminated by a mask). Observers reported displacement direction and letter identity in different blocks. Breaking object correspondence by blanking improved displacement identification but deteriorated postsaccadic letter recognition. In Experiment 2, object correspondence was broken by changing the object's contrast-polarity. There were no object displacements and observers only reported letter identity. Again, breaking object correspondence deteriorated postsaccadic letter recognition. These findings identify transsaccadic object correspondence as a key determinant of object recognition across the saccade. This is in line with the recent hypothesis that breaking object correspondence results in separate representations of presaccadic and postsaccadic objects which then compete for limited attentional processing resources (Schneider, 2013). Postsaccadic object recognition is then deteriorated because less resources are available for processing postsaccadic objects.

Keywords: saccade, visual stability, attention, object correspondence, transsaccadic memory

INTRODUCTION

Accurate vision is spatially and temporally limited. Spatially, it is limited to the fovea, the center part of the eye's retina which provides the highest visual resolution (e.g., Findlay and Gilchrist, 2003). The low resolution in the retinal periphery places a fundamental constraint on the visual exploration of the world: To view a potentially interesting object in the periphery with high acuity,

one must bring it onto the fovea by making a fast saccadic eye movement. Temporally, online visual processing is limited to fixations, discrete episodes in which the eyes stand relatively still. Every saccade interrupts useful visual input and changes the retinal position and resolution of external objects. Nevertheless, humans perceive the visual world as stable across saccades (for reviews, see Bridgeman et al., 1994; Wurtz, 2008). Moreover, coping with most natural tasks demonstrates that humans sample and relate information on task-relevant objects across eye movements (Land and Tatler, 2009; Schneider, 2013). This implies that the visual system assesses *object correspondence* across fixations (Hollingworth et al., 2008; also called object continuity, Schneider, 2013), it assesses whether input from postsaccadic and presaccadic objects (apparently) comes from the same external object (Kahneman et al., 1992; Irwin and Andrews, 1996). Object correspondence is a prerequisite for updating presaccadic low-quality information on a peripheral object with postsaccadic foveal information on the same object (Henderson and Anes, 1994; Demeyer et al., 2009; Herwig and Schneider, 2014).

Transsaccadic object correspondence and updating are considered elementary for building a task-relevant representation of the visual environment, as they tie together the samples obtained from successive fixations (Schneider, 2013; Ganmor et al., 2015; Herwig, 2015a; Wolf and Schütz, 2015; Wurtz, 2015). However, it appears that signaling of object correspondence and updating can also strikingly impair perception. An object can be displaced during a saccade for up to a third of saccade amplitude without this being noticeable (Bridgeman et al., 1975). This form of transsaccadic change-blindness suggests that the postsaccadic object location updates and overwrites the presaccadic object location (Deubel et al., 1996). As a consequence, displacement perception suffers because only the postsaccadic object location remains available (Deubel et al., 1996).

How does the visual system assess object correspondence? Object correspondence is signaled if a test of the presaccadic object against the object after the saccade results in a match (Deubel et al., 1996; Tas et al., 2012). This notion is supported by a number of studies using the *blanking paradigm*, which breaks object correspondence by blanking a saccade target object during the saccade and delaying its reappearance until shortly after eye-landing (Deubel and Schneider, 1994; Deubel et al., 1996, 1998, 2002; the discussion in terms of object correspondence comes from Tas et al., 2012). Blanking improves accuracy in reporting transsaccadic displacements of the saccade target object considerably (Deubel and Schneider, 1994; Deubel et al., 1996). In addition, blanking improves accuracy in reporting transsaccadic changes of visual object features besides location (such as spatial frequency, Weiß et al., 2015; see, also Deubel et al., 2002). Together, these results indicate that blanking prevents updating and overwriting of the presaccadic object with the postsaccadic one. Both objects are compared and this allows to identify displacements (Deubel and Schneider, 1994; Deubel et al., 1996, 2002) and changes of other visual features (Weiß et al., 2015). Briefly occluding the postsaccadic object (Deubel et al., 2002) and changing its contrast-polarity (Tas et al., 2012) helps reporting displacements in a similar way as blanking. This suggests that

breaking object correspondence in general prevents transsaccadic updating. Instead of one updated object representation, separate representations of the presaccadic and postsaccadic object should emerge (Deubel et al., 1996; Tas et al., 2012; Schneider, 2013).

Critically, the beneficial effects of breaking object correspondence for perceiving transsaccadic displacements and feature changes may come at costs in terms of postsaccadic object recognition. This hypothesis is based on the theory of "Task-dRiven visual Attention and working Memory" (TRAM, Schneider, 2013). TRAM follows the biased competition approach to attention (Desimone and Duncan, 1995) and the "Theory of Visual Attention" (Bundesen, 1990), assuming that visual objects compete for object recognition. Specifically, an object is recognized and becomes accessible (e.g., for report) if it enters capacity-limited visual working memory. An object can enter visual working memory if enough attentional processing resources (e.g., neurons, Bundesen et al., 2005) have been allocated to it. Object recognition is competitive because these processing resources are limited and have to be split among objects (Bundesen, 1990; Desimone and Duncan, 1995; Bundesen et al., 2005). Thus, the more objects take part in the competition, the less attentional processing resources are available for processing each individual object in service of object recognition. A central idea of TRAM is that the competition for object recognition is organized in discrete competition episodes of which eye fixations are a prominent case. Two kinds of objects participate in the competition. First, objects from the current episode, including those objects that have updated their corresponding counterparts from the preceding episode. Second, objects from the preceding episode for which no corresponding object was found in the current episode. Therefore, an object that has not been updated due to broken object correspondence introduces an additional competitor into the current competition episode. As a consequence, attentional processing resources must be split among more objects. This then cuts the resources for processing each individual object and thereby imposes costs on object recognition.

The present study aimed at testing the hypothesis that breaking object correspondence across the saccade deteriorates postsaccadic object recognition. Two experiments each used a different manipulation to break transsaccadic object correspondence and examined its effects on performance in a postsaccadic letter recognition task.

EXPERIMENT 1

In Experiment 1, blanking was used to break transsaccadic object correspondence (Deubel et al., 1996; cf. Tas et al., 2012). Observers made a saccade to a peripheral object which was displaced during the saccade. The postsaccadic object appeared either immediately after the saccade (no-blank condition) or after a brief blank (blank condition). A single letter was presented simultaneously to and within the postsaccadic object and was terminated by a pattern mask. Both, displacement identification and postsaccadic object recognition performance were assessed. Observers reported displacement direction and

FIGURE 1 | Paradigm of Experiment 1. Observers made a saccade to an elliptic object containing an irrelevant special character. The object was displaced during the saccade. The postsaccadic object contained a letter. It was shown for 80 ms (pattern-masked), either immediately after the saccade (no-blank condition) or after a 100 ms blank (blank condition). Displacement direction was reported in one report block, letter identity in the other. Ellipses of broken lines provide reference positions (they were not shown on the screen): A black ellipse of broken lines indicates the location of a previous object, a green ellipse of broken lines indicates the location of an upcoming object.

letter identity in two different blocks of trials. If breaking object correspondence by blanking imposes costs on object recognition, then performance in reporting the postsaccadic letter should suffer in the blank condition compared to the no-blank condition. This predicted deterioration in object recognition is diametrical to the expected performance improvement for displacement identification (Deubel and Schneider, 1994; Deubel et al., 1996, 2002).

Method
Observers
Sixteen observers (eight males, eight females) between 20 and 32 years ($Mdn = 27$ years) were paid to participate in Experiment 1. All had normal or corrected-to-normal vision (contact lenses) and gave written informed consent before the experiment. The type of experiment was approved by Bielefeld University's ethics committee.

Apparatus and Stimuli
The experiment took place in a dimly lit room. Eye behavior was recorded by a video-based tower-mounted eye-tracker (Eyelink 1000, SR Research, Mississauga, ON, Canada) which was calibrated using a nine-point grid procedure and sampled observers' right eyes at 1000 Hz. Observers' heads were stabilized by forehead and chin rests, 71 cm from the 19"-CRT-screen (G90FB, ViewSonic, Brea, CA, USA) which ran with a resolution of 1024 × 768 pixels (at physical dimensions of 36 cm × 27 cm) and a refresh rate of 100 Hz.

The experiment was controlled by Experiment Builder (v1.10.1025). Stimulus luminance was measured using a MAVOLUX-digital luminance meter (Gossen, Nuremberg, Germany). Stimuli were black (<1 cd $\times$ m^{-2}) special characters ($\$\$\&\}/[\mu\sim$) and letters (ABDGHJKLMNPRSTVX;

$0.48° \times 0.56°$) in Arial font and a black plus-character ($0.28° \times 0.28°$) was used as fixation cross. The saccade target object was a gray ellipse (29 cd $\times$ m^{-2}; $0.7° \times 1.26°$). The white background had a luminance of 89 cd $\times$ m^{-2}. Four different pattern masks were used, which consisted of rectangles ($1.01° \times 1.5°$) filled with black scrambled lines of different widths.

Design and Procedure
Figure 1 illustrates the experimental paradigm. Each trial began with fixation of a central fixation cross (at least 490 ms continuous fixation plus a variable delay between 0 and 500 ms; trials were aborted and repeated if the fixation cross was not fixated). Afterward, the fixation cross was extinguished and an ellipse was shown as saccade target object, 6° or 8° from screen center in horizontal direction. This ellipse contained an irrelevant special character and was presented until the observer made a saccade to it (detected using velocity and acceleration thresholds of $30° \times s^{-1}$ and $8000° \times s^{-2}$). In the no-blank condition, the now empty ellipse was displaced for 1° during the saccade (with the next screen refresh after saccade detection). Initial position of the ellipse (6° or 8°, left or right to screen center) and displacement direction (left or right) were randomized across trials with equal numbers of occurrence in each condition. At the next screen refresh after eye-landing, a letter was shown within the ellipse for 80 ms and terminated by a pattern mask lasting for 300 ms. The letter was randomly drawn from the set of used letters (each letter occurred equally often in each blanking condition and report block; special characters were drawn analogously). The mask was drawn randomly from the set of used masks. After 500 ms, a response screen prompted observers to report letter identity or displacement direction using the keyboard (unspeeded forced choice; letter-keys or "F1" and "F12", respectively). The next
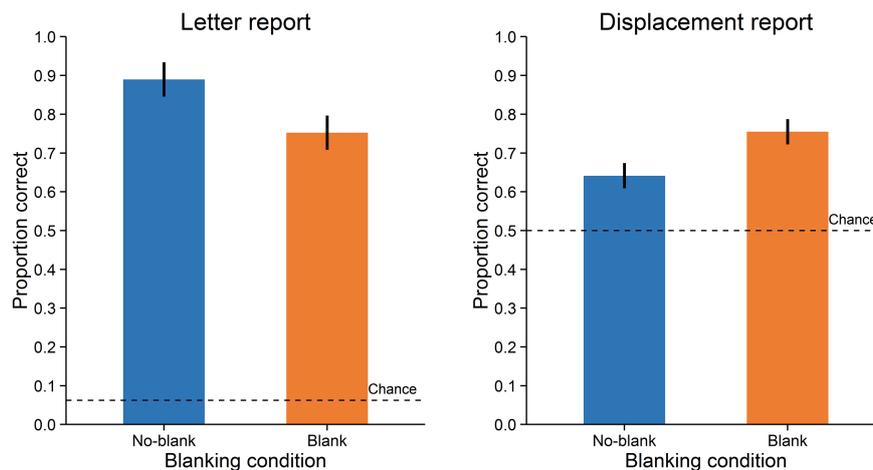
FIGURE 2 | Performance in Experiment 1. Letter report performance **(left)** and displacement report performance **(right)**. Error-bars indicate 95% confidence intervals for within-subjects designs (Morey, 2008). Broken lines indicate chance level.

trial started after the report was made. The blank condition was identical to the no-blank condition except that an empty screen was shown during the saccade and lasted for another 100 ms from the screen refresh after the eye-landing. Trials of the two blanking conditions occurred in random order within report blocks. All observers performed two report blocks (order counterbalanced across sample) of 152 trials, the half of which belonging to the no-blank and the other to the blank condition. In these blocks, they either only reported displacement direction or only letter identity. Observers performed 16 training trials before each report block.

Results

Trials were excluded from the analyses, if no saccade was made until 400 ms after onset of the saccade target object, saccade latency was below 100 ms (anticipatory saccades), or the saccade target object was missed by more than 2.5°. A total of 4.3% of the trials was discarded. Letter and displacement reports were each pooled across trials on which saccade target objects appeared 6° or 8° to the left or right of fixation (Deubel et al., 1996). They were also pooled across orders of displacement and letter report blocks because mixed analyses of variances (ANOVAs) showed that neither order nor the interaction of order and blanking conditions affected letter or displacement report performance, all $Fs < 3.167$, all $ps > 0.096$.

Accuracy was assessed as the proportion of correct responses. A paired-samples t-test with d_z (Cohen, 1988) as effect size showed that letter reports were significantly more accurate in the no-blank ($M = 0.89$, $SD = 0.11$) compared to the blank condition ($M = 0.75$, $SD = 0.17$), $t(15) = 4.671$, $p < 0.001$, $d_z = 1.17$, Bayes Factor (BF) = 108.271, (**Figure 2**, left; Bayes Factors were computed using the BayesFactor (0.9.10-2) package for R (3.0.3), cf. Rouder et al., 2009, values greater one support the alternative and values smaller one the null hypothesis). In contrast, displacement reports were significantly less accurate in the no-blank ($M = 0.64$, $SD = 0.12$) than in the blank
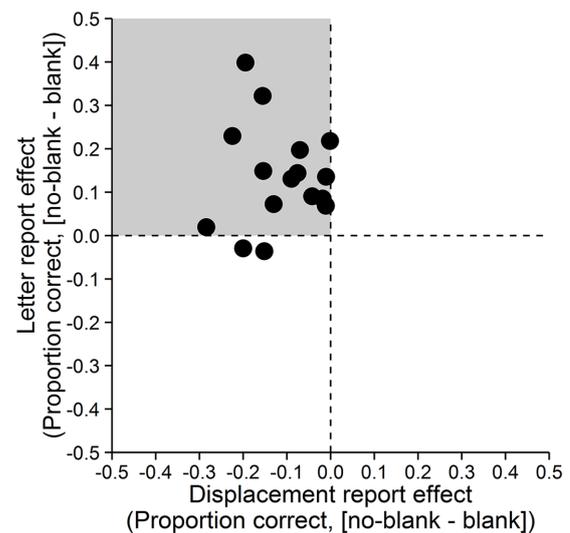

FIGURE 3 | Effects of blanking on letter and displacement reports for individual observers. Differences between the no-blank and blank condition for both, displacement report (x-axis) and letter report (y-axis). Each point represents the value of one observer. The gray quadrant indicates the region in which points should fall if the effect of blanking on displacement report performance is in the opposite direction of the effect of blanking on letter report performance.

condition ($M = 0.75$, $SD = 0.16$), $t(15) = -5.238$, $p < 0.001$, $d_z = -1.31$, $BF = 284.724$, (**Figure 2**, right). As evident from **Figure 3**, the effects of blanking on letter report performance and on displacement report performance were in opposite direction for most observers.

Not surprisingly, observers' mean saccade latencies (i.e., the time between the onset of the saccade target object and saccade detection) did not differ significantly between the blanking conditions, both in the letter report block (no-blank: $M = 132$ ms,

$SD = 11$ ms, blank: $M = 133$ ms, $SD = 10$ ms), $t(15) = -1.756$, $p = 0.100$, $d_z = -0.44$, $BF = 0.893$ and in the displacement report block (no-blank: $M = 166$ ms, $SD = 21$ ms, blank: $M = 168$, $SD = 22$ ms), $t(15) = -0.858$, $p = 0.404$, $d_z = -0.21$, $BF = 0.352$. The blanking conditions did not significantly differ in deviations of gaze positions from the postsaccadic object in the eye tracker's first sample after the onset of the postsaccadic object (observers' mean distance between gaze position and postsaccadic object), neither in the letter report block (no-blank: $M = 1.14°$, $SD = 0.13°$, blank: $M = 1.18°$, $SD = 0.13°$), $t(15) = -1.730$, $p = 0.104$, $d_z = -0.43$, $BF = 0.864$, nor in the displacement report block (no-blank: $M = 1.16°$, $SD = 0.14°$, blank: $M = 1.18°$, $SD = 0.14°$), $t(15) = -0.545$, $p = 0.594$, $d_z = -0.14$, $BF = 0.291$. Likewise, the blanking conditions did not significantly differ in variability of gaze positions in these samples of the eye tracker (observers' standard deviation of distances between gaze position and postsaccadic object), neither in the letter report block (no-blank: $M = 0.45°$, $SD = 0.09°$, blank: $M = 0.47°$, $SD = 0.08°$), $t(15) = -1.397$, $p = 0.183$, $d_z = -0.35$, $BF = 0.579$, nor in the displacement report block (no-blank: $M = 0.53°$, $SD = 0.11°$, blank: $M = 0.54°$, $SD = 0.10°$), $t(15) = -0.437$, $p = 0.669$, $d_z = -0.11$, $BF = 0.278$.

Discussion

Experiment 1 provides first support for the hypothesis that breaking object correspondence across the saccade impairs postsaccadic object recognition (Schneider, 2013). Recognition of a postsaccadic letter was deteriorated in the blank condition, where object correspondence was broken, compared to the no-blank condition, where it was not broken. In stark contrast, breaking object correspondence by blanking was beneficial for identifying transsaccadic object displacements. This beneficial effect of blanking on perception of transsaccadic object displacements replicates previous work and shows that the present blanking manipulation was effective (Deubel and Schneider, 1994; Deubel et al., 1996, 2002).

It is well-established that blanking breaks transsaccadic object correspondence (Tas et al., 2012) and prevents the updating and overwriting of presaccadic object information (Deubel and Schneider, 1994; Deubel et al., 1996, 2002; Weiß et al., 2015). However, some issues must be considered before we can conclude that the present deterioration in postsaccadic letter recognition was in fact due to broken object correspondence. First, the deterioration might have been due to the different temporal intervals between eye-landing and onset of the postsaccadic object in the two blanking conditions. Visual processing has been claimed to be enhanced immediately after saccades (Ibbotson and Krekelberg, 2011). Thus, processing of the postsaccadic letter might have been enhanced when the object was immediately visible after the saccade in the no-blank condition compared to when it appeared later in the blank condition. Second, the onset of the postsaccadic object was visible in the blank condition but was concealed by the saccade in the no-blank condition (e.g., Krock and Moore, 2015). Therefore, the deterioration might also stem from interference of this onset with recognition of the letter (as a form of masking; e.g., Enns and Di Lollo, 2000). Third, objects were

always displaced during the saccade and this may have affected postsaccadic object recognition differently in the two blanking conditions. In line with these alternative explanations, one might suppose that object correspondence was broken in both blanking conditions, meaning it cannot account for the deteriorated postsaccadic letter recognition. This might have been the case because in both conditions a special character in the presaccadic object changed into a letter in the postsaccadic object (cf. Demeyer et al., 2010). To rule out these alternative explanations, Experiment 2 examined how postsaccadic letter recognition was affected by manipulating transsaccadic object correspondence in conditions with identical time courses and without any object displacements.

EXPERIMENT 2

In Experiment 2, a change of contrast-polarity was used to break transsaccadic object correspondence (Tas et al., 2012). Observers made a saccade to a peripheral object which was black or white. The contrast-polarity of this object either stayed the same (no-change condition) or changed during the saccade (change condition) so that a black presaccadic object changed into a white postsaccadic one and vice versa. Similar to Experiment 1, a single letter appeared simultaneously to and within the postsaccadic object and was terminated by a pattern mask. In contrast to Experiment 1, however, both of these polarity-change conditions were identical in time course and there were no intrasaccadic object displacements. Observers' only task was to report the postsaccadic letter. Now, if breaking object correspondence by changing contrast-polarity imposes costs on postsaccadic object recognition, then performance in reporting the postsaccadic letter should suffer in the change compared to the no-change condition.
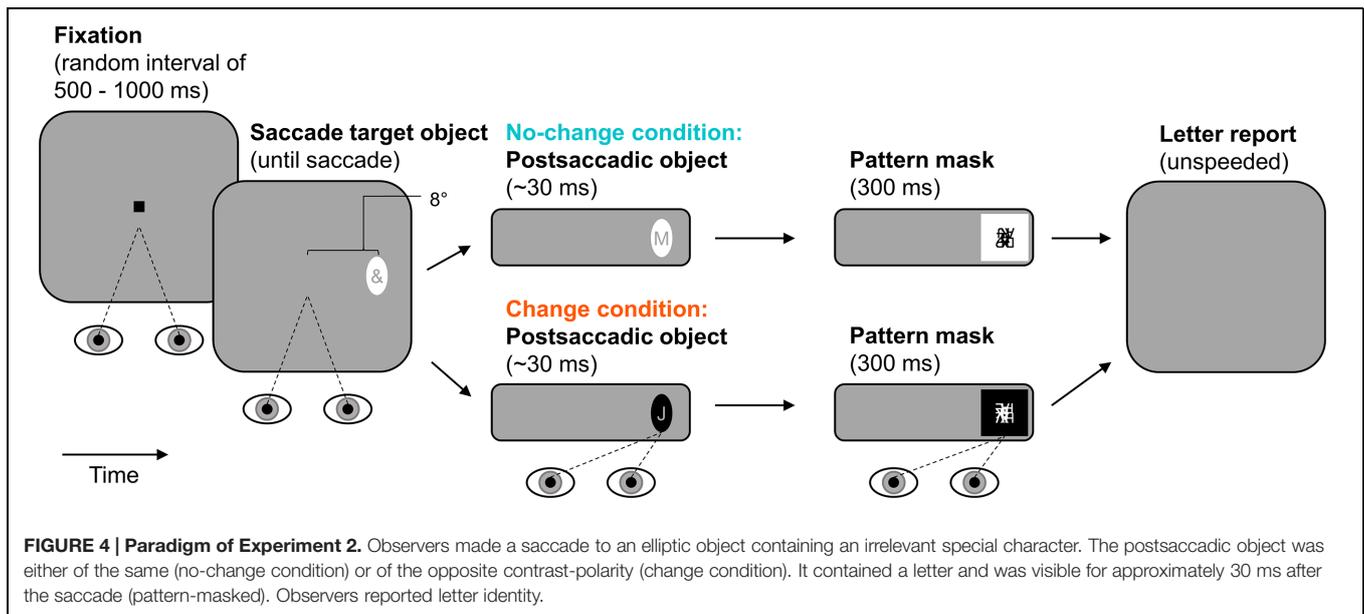
Method
Observers

Twelve observers (2 males, 10 females) were paid to take part in Experiment 2. They were between 21 and 31 years old ($Mdn = 27$), all had normal or corrected-to-normal vision (contact lenses) and gave written informed consent before the experiment. The type of experiment was approved by Bielefeld University's ethics committee.

Apparatus and Stimuli

The apparatus and testing conditions in Experiments 1 and 2 were identical but not the same (i.e., the monitors were of the same model but were two different ones). Besides, a desktop-mounted video-based eye-tracker (Eyelink 1000, SR Research, Mississauga, ON, Canada) recorded eye behavior in Experiment 2.

Experiment 2 was controlled by the Psychophysics Toolbox (3.0.12; Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) and Eyelink Toolbox (3.0.12; Cornelissen et al., 2002) extensions for MATLAB R2014b (The MathWorks, Natick, MA, USA). Stimuli were gray (67 cd $\times$ m^{-2}) special characters (%#$&; $0.4° \times 0.4°$) and letters (ABDEFGHJKLMNOPRSTVXZ; $0.32° \times 0.4°$) in

FIGURE 4 | Paradigm of Experiment 2. Observers made a saccade to an elliptic object containing an irrelevant special character. The postsaccadic object was either of the same (no-change condition) or of the opposite contrast-polarity (change condition). It contained a letter and was visible for approximately 30 ms after the saccade (pattern-masked). Observers reported letter identity.

Arial font and saccade target objects were black (1 cd $\times$ m^{-2}) or white (135 cd $\times$ m^{-2}) ellipses (0.65° $\times$ 1.05°). The gray background had a luminance of 67 cd $\times$ m^{-2}. A black square (0.1° $\times$ 0.1°) was used as central fixation stimulus. Ninety-nine pattern masks were algorithmically created for each observer and for both, black and white ellipses. This relatively large number of masks was chosen to minimize adaptation to the masks. The masks consisted of black or white rectangles (2° $\times$ 2°), each containing nine letters that were drawn randomly without replacement from the set of used letters. These letters were mirror-reversed and upside-down, they overlapped partially, and together covered an area of about 1° $\times$ 1° within a rectangle. For black rectangles the letters were white and for white rectangles they were black.

Design and Procedure

The experimental paradigm is illustrated in **Figure 4**. Observers started each trial by pressing the space-bar. In the beginning of a trial, observers fixated a central fixation stimulus for a random interval ranging from 500 to 1000 ms. Afterward, the fixation stimulus was extinguished and an ellipse was presented as saccade target object 8° to the left or right of screen center (randomized across trials with equal numbers of occurrence in each condition). The ellipse contained an irrelevant special character (randomly drawn from the set of used special characters) and stayed on screen until the observer made a saccade to it (detected using velocity and acceleration thresholds of 35° $\times$ s^{-1} and 9500° $\times$ s^{-2}). This presaccadic ellipse was either black or white. The postsaccadic ellipse contained a letter (randomly drawn from the set of used letters) and appeared during the saccade, that is, on the next screen refresh after detection of saccade onset. In the no-change condition, the postsaccadic ellipse and the presaccadic ellipse were identical in their contrast-polarity. In the change condition, the postsaccadic ellipse was of the opposite contrast-polarity of the presaccadic ellipse. That is, a black presaccadic
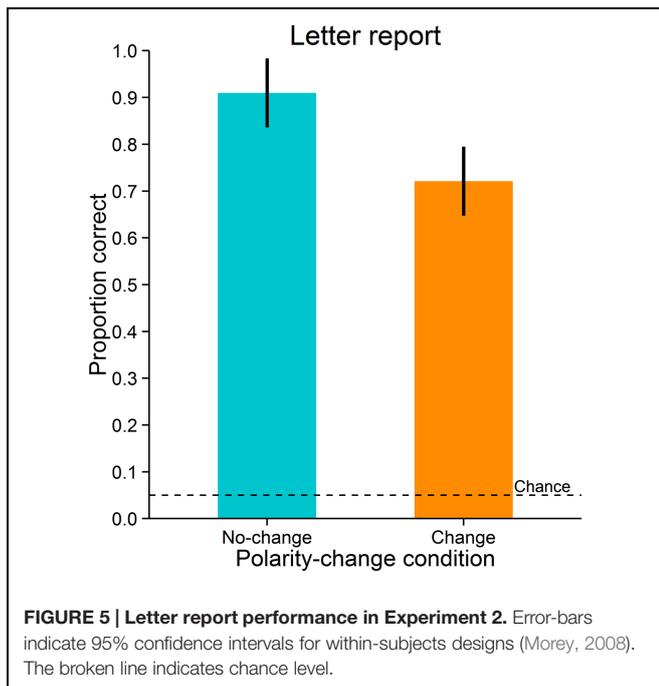
ellipse changed into a white postsaccadic one and vice versa. Whether presaccadic ellipses were black or white was randomized across trials but the number of occurrences was equal in the two polarity-change conditions. The postsaccadic ellipse was followed by a pattern mask of the same polarity. The mask was presented two or three screen refreshes after detection of saccade end so that the postsaccadic ellipse was visible after the saccade for 31 ms on average ($SD = 3$ ms). The mask was drawn randomly from the set of created masks and lasted for 300 ms. After that, the screen went blank and observers reported the letter using the keyboard (unspeeded forced-choice). They could start the next trial after 100 ms.

Observers performed 64 trials of each polarity-change condition in randomized order. Trials were aborted and repeated on a randomly chosen subsequent trial if observers failed to fixate the central fixation stimulus or if they missed the saccade target object by more than 2.5°. In this way, a total of 22.5% of the trials was repeated. Observers performed 32 training trials before the experiment.

Results

Seven trials were excluded from analysis because saccade latency was below 100 ms or above 400 ms. Letter reports were pooled across trials on which saccade target objects appeared to the left or right of screen center (as for Experiment 1). They were also pooled across trials with different presaccadic ellipse polarities because a repeated-measures ANOVA indicated that neither presaccadic ellipse polarity nor its interaction with the two polarity-change conditions (i.e., no-change or change) affected letter report performance, both Fs < 0.099, both ps > 0.758 (although distributions of proportions of correct responses were negatively skewed for both presaccadic ellipse polarities in the no-change condition).

Accuracy was measured as the proportion of correct responses. Letter reports were significantly more accurate in

FIGURE 5 | Letter report performance in Experiment 2. Error-bars indicate 95% confidence intervals for within-subjects designs (Morey, 2008). The broken line indicates chance level.

the no-change condition ($M = 0.91$, $SD = 0.15$) than in the change condition ($M = 0.72$, $SD = 0.20$), $t(11) = 3.989$, $p = 0.002$, $d_z = 1.15$; $BF = 21.223$ (**Figure 5**). As can be expected, the two polarity-change conditions did not differ significantly in observers' mean saccade latencies (no-change condition: $M = 155$ ms, $SD = 20$ ms; change condition: $M = 155$ ms, $SD = 21$ ms), $t(11) = -0.494$, $p = 0.631$, $d_z = -0.14$, $BF = 0.319$. Likewise, the conditions did not differ significantly in deviations of saccade landing positions from saccade target objects (observers' mean distances between saccade landing positions and saccade target objects; no-change condition: $M = 0.77°$, $SD = 0.19°$; change condition: $M = 0.79°$, $SD = 0.19°$), $t(11) = -1.665$, $p = 0.124$, $d_z = -0.48$, $BF = 0.846$. Also, they did not differ significantly in variability of deviations of saccade landing positions from saccade target objects (observers' standard deviations of distances between saccade landing positions and saccade target objects; no-change condition: $M = 0.35°$, $SD = 0.07°$; change condition: $M = 0.35°$, $SD = 0.07°$), $t(11) = -0.216$, $p = 0.833$, $d_z = -0.06$, $BF = 0.293$.

Discussion

Experiment 2 provides further evidence that breaking transsaccadic object correspondence impairs postsaccadic object recognition (Schneider, 2013). Recognition of a postsaccadic letter was deteriorated in the change condition, where object correspondence was broken, compared with the no-change condition, where it was not broken. As such, the findings of Experiment 2 perfectly replicate and extend the findings from Experiment 1. Moreover, Experiment 2 also controlled for alternative interpretations of the findings of Experiment 1.

In Experiment 2, transsaccadic object correspondence was broken by changing contrast-polarity rather than by blanking. This allowed to keep the temporal interval between eye-landing and onset of the postsaccadic object constant in the two polarity-change conditions. Therefore, in contrast to Experiment 1, there were no differences in time course between conditions which could account for the differences in postsaccadic letter recognition. For this reason, two alternative explanations of the findings of Experiment 1 can be dismissed for the ones of Experiment 2. First, the differences in postsaccadic letter recognition did not result from enhanced processing immediately after saccades (Ibbotson and Krekelberg, 2011), because letter recognition would have been enhanced in both polarity-change conditions. Second, the differences did not result from interference of the onset of the postsaccadic object with letter recognition, because this onset happened during the saccade and likewise in both polarity-change conditions. Furthermore and again contrasting Experiment 1, there were no object displacements in Experiment 2. This excludes any differential effects of displacements between conditions. Both experiments had in common, however, that the presaccadic object contained an irrelevant special character which changed into a letter in the postsaccadic object. Although this change might have broken object correspondence (Demeyer et al., 2010), this cannot refer to the results of Experiment 2. The character change occurred in both polarity-change conditions and notwithstanding there was a pronounced effect of the polarity change on postsaccadic letter recognition. It has been shown previously that changing contrast-polarity is an effective tool to break transsaccadic object correspondence (Tas et al., 2012). Thus, even if the effect of changing contrast-polarity only added to the effect of changing the special character into the letter, it still demonstrates an effect of object correspondence on object recognition. Taken together, the findings of Experiment 2 therefore strongly argue that breaking object correspondence across the saccade deteriorates postsaccadic object recognition.

GENERAL DISCUSSION

We asked whether breaking object correspondence across the saccade impairs postsaccadic object recognition. The present findings indicate that this is the case. In both of our experiments, recognition of a postsaccadic letter was deteriorated when transsaccadic object correspondence was broken, compared with when it was not broken. Now we can ask which cognitive mechanisms might underlie these effects.

One possible interpretation of the present findings is that breaking transsaccadic object correspondence increases locational uncertainty of task-relevant information after the saccade. The precision of saccades is limited so that there is always variation in saccade landing positions. Therefore, to sample information on a saccade target object after a saccade, this object must be re-located (Hollingworth et al., 2008), even if it remained at its location across the saccade.

Breaking transsaccadic object correspondence may hinder this re-location (and this might already happen during the saccade, Panouillères et al., 2013). Information on where to find task-relevant information after the saccade would then be less specific. This could impair postsaccadic object recognition, for instance because less attentional processing resources would be devoted to the location of the postsaccadic object.

Alternatively, intact transsaccadic object correspondence may provide computational savings which are lost in case object correspondence is broken. Specifically, new high-resolution foveal information on a postsaccadic object updates the representation of the corresponding presaccadic object (Tas et al., 2012; cf. Deubel and Schneider, 1994; Deubel et al., 1996, 2002). In contrast, if transsaccadic object correspondence is broken, then there is no presaccadic representation that can (or should) be updated with postsaccadic information. An entirely new representation must be created for the postsaccadic object. This additional requirement may delay processing of the postsaccadic object (such delays have for instance been found when monkeys had to adapt their smooth pursuit eye movements to postsaccadic motion patterns, Fallah and Reynolds, 2012). Such processing delays then deteriorate the postsaccadic recognition of objects and this is most prominent when postsaccadic objects are only briefly available (as in the current experiments).

These two interpretations suggest a close link between transsaccadic object correspondence and postsaccadic object recognition. However, they do not provide a mechanistic theory of the relationship between these processes. In contrast, TRAM (Schneider, 2013) may deliver a first step toward such a theory by proposing that *attentional weights* (Bundesen, 1990) are not only mediating competition for access to visual working memory across saccades but that they should also establish correspondence between presaccadic and postsaccadic objects.

Attentional weights represent the processing priority of objects by combining the task-driven and the intrinsic relevance of object features (Bundesen, 1990). Neuronally, attentional weights are assumed to exist in spatially organized priority maps in several brain areas (Bundesen et al., 2005; cf. Fecteau and Munoz, 2006; Cavanagh et al., 2010; Zelinsky and Bisley, 2015). Thus, attentional weights code for the feature-derived attentional priority of objects but also for their spatial location. With this combination of priority and location, attentional weights can provide a number of functions fundamental for human active vision. Within priority maps, attentional weights control saccade target selection ("where-to-look-next?", Wischnewski et al., 2009, 2010; Schneider, 2013). This is a form of selection-for-action (Allport, 1987; Neumann, 1987). In addition, attentional weights govern the allocation of neuronal processing resources to objects in order to accomplish object recognition (Bundesen et al., 2005). This is selection-for-perception (covert visual attention). Selection-for-action and selection-for-perception are assumed to be tightly coupled (Schneider, 1995; Schneider and Deubel, 2002; cf. Irwin and Gordon, 1998) and attentional weights in priority maps may establish this coupling (Schneider,

2013; Herwig, 2015b). Furthermore, attentional weights (in this context called "attentional pointers") can align presaccadic and postsaccadic information by keeping track of object locations across saccades (Cavanagh et al., 2010). This proposal is based on studies showing that the location sensitivity of neurons in some priority maps (i.e., the maps assumed to implement attentional weights, cf. Bundesen et al., 2005) is updated before saccades to accommodate impending saccade-induced changes of retinal locations (Duhamel et al., 1992). Along these lines, TRAM proposes that the attentional weight of a presaccadic and a postsaccadic object is used to test for object correspondence across saccades (Schneider, 2013). Object correspondence is then signaled if the attentional weight of the postsaccadic object matches the attentional weight that is predicted based on the presaccadic object. Thereby, the attentional weight could spatially route postsaccadic feature input to presaccadically created object representations in the process of transsaccadic updating. This may give rise to visual stability: the perception of a stable world despite the retinal image changes induced by saccades (e.g., Mathôt and Theeuwes, 2011).

In contrast, if object correspondence is broken, the visual system signals that a new object has appeared after the saccade (Kahneman et al., 1992; Irwin and Andrews, 1996). According to TRAM, the attentional weight of the presaccadic object is then encapsulated (i.e., retained with its current connection to presaccadic features) to protect the presaccadic object against being updated and overwritten by the new (non-corresponding) postsaccadic object. This encapsulated attentional weight competes with the attentional weights of postsaccadic objects. Neuronal processing resources are normalized over all present attentional weights (e.g., Bundesen et al., 2005; Poth et al., 2014). Instead of having all neuronal resources available for processing objects of the postsaccadic competition episode, some amount of resources is again (Schneider, 2013) or still (Petersen et al., 2012) allocated to the presaccadic object. In sum, TRAM proposes that breaking object correspondence across the saccade provokes attentional competition between presaccadic and postsaccadic objects. This attentional competition hypothesis provides one explanation why breaking object correspondence impaired postsaccadic object recognition in the present experiments. Testing the hypothesis may be an interesting avenue for future studies aiming to bridge research on transsaccadic object correspondence and on mechanisms of visual attention and object recognition.

CONCLUSION

The present study shows for the first time that breaking object correspondence across the saccade deteriorates postsaccadic object recognition. This reveals a crucial role of object correspondence for vision across successive fixations and saccades. Natural human vision consists of a succession of fixations and saccadic eye movements. Therefore, classical theories of task-driven object recognition (and visual attention; Bundesen, 1990; Wolfe, 1994) should now take mechanisms of transsaccadic object correspondence into account.

AUTHOR CONTRIBUTIONS

CP, AH, and WS designed and planned the research. CP programmed the experiments and analyzed the data. CP, AH, and WS wrote the paper.

FUNDING

ACKNOWLEDGMENT

REFERENCES

Allport, A. (1987). "Selection for action: Some behavioral and neurophysiological considerations of attention and action," in *Perspectives on Perception and Action*, eds H. Heuer and A. Sanders (Hillsdale, NJ: Lawrence Erlbaum Associates), 395–419.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357

Bridgeman, B., Hendry, D., and Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vis. Res.* 15, 719–722. doi: 10.1016/0042-6989(75)90290-4

Bridgeman, B., van der Heijden, A. H. C., and Velichkovsky, B. M. (1994). A theory of visual stability across saccadic eye movements. *Behav. Brain Sci.* 17, 247–292. doi: 10.1017/S0140525X00034361

Bundesen, C. (1990). A theory of visual attention. *Psychol. Rev.* 97, 523–547. doi: 10.1037/0033-295X.97.4.523

Bundesen, C., Habekost, T., and Kyllingsbæk, S. (2005). A neural theory of visual attention: bridging cognition and neurophysiology. *Psychol. Rev.* 112, 291–328. doi: 10.1037/0033-295X.112.2.291

Cavanagh, P., Hunt, A. R., Afraz, A., and Rolfs, M. (2010). Visual stability based on remapping of attention pointers. *Trends Cogn. Sci.* 14, 147–153. doi: 10.1016/j.tics.2010.01.007

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. New York, NY: Psychology Press.

Cornelissen, F. W., Peters, E. M., and Palmer, J. (2002). The eyelink toolbox: eye tracking with MATLAB and the psychophysics toolbox. *Behav. Res. Methods* 34, 613–617. doi: 10.3758/BF03195489

Demeyer, M., De Graef, P., Wagemans, J., and Verfaillie, K. (2009). Transsaccadic identification of highly similar artificial shapes. *J. Vis.* 9, 1–14. doi: 10.1167/9.4.28

Demeyer, M., De Graef, P., Wagemans, J., and Verfaillie, K. (2010). Object form discontinuity facilitates displacement discrimination across saccades. *J. Vis.* 10, 1–14. doi: 10.1167/10.6.17

Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222. doi: 10.1146/annurev.ne.18.030195.001205

Deubel, H., Bridgeman, B., and Schneider, W. X. (1998). Immediate postsaccadic information mediates space constancy. *Vis. Res.* 38, 3147–3159. doi: 10.1016/S0042-6989(98)00048-0

Deubel, H., and Schneider, W. X. (1994). Perceptual stability and postsaccadic visual information: can man bridge a gap? *Behav. Brain Sci.* 17, 259–260. doi: 10.1017/S0140525X00034397

Deubel, H., Schneider, W. X., and Bridgeman, B. (1996). Postsaccadic target blanking prevents saccadic suppression of image displacement. *Vis. Res.* 36, 985–996. doi: 10.1016/0042-6989(95)00203-0

Deubel, H., Schneider, W. X., and Bridgeman, B. (2002). "Transsaccadic memory of position and form," in *Progress in Brain Research. The Brain's Eye: Neurobiological and Clinical Aspects of Oculomotor Research*, eds J. Hyona, D. P. Munoz, W. Heide, and R. Radach (Amsterdam: Elsevier), 165–180. doi: 10.1016/S0079-6123(02)40049-0

Duhamel, J.-R., Colby, C. L., and Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* 255, 90–92. doi: 10.1126/science.1553535

Enns, J. T., and Di Lollo, V. (2000). What's new in visual masking? *Trends Cogn. Sci.* 4, 345–352. doi: 10.1016/S1364-6613(00)01520-5

Fallah, M., and Reynolds, J. H. (2012). Contrast dependence of smooth pursuit eye movements following a saccade to superimposed targets. *PLoS ONE* 7:e37888. doi: 10.1371/journal.pone.0037888

Fecteau, J. H., and Munoz, D. P. (2006). Salience, relevance, and firing: a priority map for target selection. *Trends Cogn. Sci.* 10, 382–390. doi: 10.1016/j.tics.2006.06.011

Findlay, J. M., and Gilchrist, I. D. (2003). *Active Vision: The Psychology of Looking and Seeing*. Oxford: Oxford University Press.

Ganmor, E., Landy, M. S., and Simoncelli, E. P. (2015). A near-optimal integration of orientation information across saccades. *J. Vis.* 15, 1–17. doi: 10.1167/15.16.8

Henderson, J. M., and Anes, M. D. (1994). Roles of object-file review and type priming in visual identification within and across eye fixations. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 826–839. doi: 10.1037/0096-1523.20.4.826

Herwig, A. (2015a). Transsaccadic integration and perceptual continuity. *J. Vis.* 15, 1–6. doi: 10.1167/15.16.7

Herwig, A. (2015b). Linking perception and action by structure or process? Toward an integrative perspective. *Neurosci. Biobehav. Rev.* 52, 105–116. doi: 10.1016/j.neubiorev.2015.02.013

Herwig, A., and Schneider, W. X. (2014). Predicting object features across saccades: evidence from object recognition and visual search. *J. Exp. Psychol. Gen.* 143, 1903–1922. doi: 10.1037/a0036781

Hollingworth, A., Richard, A. M., and Luck, S. J. (2008). Understanding the function of visual short-term memory: transsaccadic memory, object correspondence, and gaze correction. *J. Exp. Psychol.* 137, 163–181. doi: 10.1037/0096-3445.137.1.163

Ibbotson, M., and Krekelberg, B. (2011). Visual perception and saccadic eye movements. *Curr. Opin. Neurobiol.* 21, 553–558. doi: 10.1016/j.conb.2011.05.012

Irwin, D. E., and Andrews, R. V. (1996). "Integration and accumulation of information across saccadic eye movements," in *Attention and Performance XVI: Information Integration in Perception and Action*, eds I. Toshio and J. L. McClelland (Oxford: Oxford University Press), 125–155.

Irwin, D. E., and Gordon, R. D. (1998). Eye movements, attention, and trans-saccadic memory. *Vis. Cogn.* 5, 127–155. doi: 10.1080/713756783

Kahneman, D., Treisman, A., and Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cogn. Psychol.* 24, 175–219. doi: 10.1016/0010-0285(92)90007-O

Kleiner, M., Brainard, D., and Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception* 36(ECVP Abstr. Suppl.).

Krock, R. M., and Moore, T. (2015). The influence of gaze control on visual perception: eye movements and visual stability. *Cold Spring Harb. Symp. Quant. Biol.* 79, 123–130. doi: 10.1101/sqb.2014.79.024836

Land, M. F., and Tatler, B. W. (2009). *Looking and Acting: Vision and Eye Movements in Natural Behaviour*. Oxford: Oxford University Press.

Mathôt, S., and Theeuwes, J. (2011). Visual attention and stability. *Philos. Trans. R. Soc. B Biol. Sci.* 366, 516–527. doi: 10.1098/rstb.2010.0187

Morey, R. D. (2008). Confidence intervals from normalized data: a correction to Cousineau (2005). *Tutorial Quant. Methods Psychol.* 4, 61–64.

Neumann, O. (1987). "Beyond capacity: a functional view of attention," in *Perspectives on Perception and Action*, eds H. Heuer and A. Sanders (Hillsdale, NJ: Lawrence Erlbaum Associates), 361–394.

Panouillères, M., Gaveau, V., Socasau, C., Urquizar, C., and Pélisson, D. (2013). Brain processing of visual information during fast eye movements maintains motor performance. *PLoS ONE* 8:e54641. doi: 10.1371/journal.pone.0054641

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897X00366

Petersen, A., Kyllingsbæk, S., and Bundesen, C. (2012). Measuring and modeling attentional dwell time. *Psychon. Bull. Rev.* 19, 1029–1046. doi: 10.3758/s13423-012-0286-y

Poth, C. H., Petersen, A., Bundesen, C., and Schneider, W. X. (2014). Effects of monitoring for visual events on distinct components of attention. *Front. Psychol.* 5:930. doi: 10.3389/fpsyg.2014.00930

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., and Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* 16, 225–237. doi: 10.3758/PBR.16.2.225

Schneider, W. X. (1995). VAM: a neuro-cognitive model for visual attention control of segmentation, object recognition and space-based motor action. *Vis. Cogn.* 2, 331–375. doi: 10.1080/13506289508401737

Schneider, W. X. (2013). Selective visual processing across competition episodes: a theory of task-driven visual attention and working memory. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368:20130060. doi: 10.1098/rstb.2013.0060

Schneider, W. X., and Deubel, H. (2002). "Selection-for-perception and selection-for-spatial-motor-action are coupled by visual attention: a review of recent findings and new evidence from stimulus-driven saccade control," in *Attention and Performance XIX: Common Mechanisms in Perception and Action*, eds W. Prinz and B. Hommel (Oxford: Oxford University Press), 609–627.

Tas, A. C., Moore, C. M., and Hollingworth, A. (2012). An object-mediated updating account of insensitivity to transsaccadic change. *J. Vis.* 12, 1–13. doi: 10.1167/12.11.18

Weiß, K., Schneider, W. X., and Herwig, A. (2015). A "blanking effect for surface features: transsaccadic spatial-frequency discrimination is improved by postsaccadic blanking. *Attent. Percept. Psychophys.* 77, 1500–1506. doi: 10.3758/s13414-015-0926-1

Wischnewski, M., Belardinelli, A., Schneider, W. X., and Steil, J. J. (2010). Where to look next? Combining static and dynamic proto-objects in a TVA-based model of visual attention. *Cogn. Comput.* 2, 326–343. doi: 10.1007/s12559-010-9080-1

Wischnewski, M., Steil, J. J., Kehrer, L., and Schneider, W. X. (2009). Integrating inhomogeneous processing and proto-object formation in a computational model of visual attention. *Cogn. Syst. Monogr.* 6, 93–102. doi: 10.1007/978-3-642-10403-9_10

Wolf, C., and Schütz, A. C. (2015). Trans-saccadic integration of peripheral and foveal feature information is close to optimal. *J. Vis.* 15, 1–18. doi: 10.1167/15.16.1

Wolfe, J. M. (1994). Guided search 2.0: a revised model of visual search. *Psychon. Bull. Rev.* 1, 202–238. doi: 10.3758/BF03200774

Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vis. Res.* 48, 2070–2089. doi: 10.1016/j.visres.2008.03.021

Wurtz, R. H. (2015). Brain mechanisms for active Vision. *Daedalus* 144, 10–21. doi: 10.1162/DAED_a_00314

Zelinsky, G. J., and Bisley, J. W. (2015). The what, where, and why of priority maps and their interactions with visual working memory. *Ann. N. Y. Acad. Sci.* 1339, 154–164. doi: 10.1111/nyas.12606

Study 2

Breaking object correspondence across saccades impairs object recognition: The role of color and luminance

Breaking object correspondence across saccades impairs object recognition: The role of color and luminance

Christian H. Poth

Neuro-cognitive Psychology, Department of Psychology and Cluster of Excellence Cognitive Interaction Technology, Bielefeld University, Bielefeld, Germany

Werner X. Schneider

Neuro-cognitive Psychology, Department of Psychology and Cluster of Excellence Cognitive Interaction Technology, Bielefeld University, Bielefeld, Germany

Rapid saccadic eye movements bring the foveal region of the eye's retina onto objects for high-acuity vision. Saccades change the location and resolution of objects' retinal images. To perceive objects as visually stable across saccades, correspondence between the objects before and after the saccade must be established. We have previously shown that breaking object correspondence across the saccade causes a decrement in object recognition (Poth, Herwig, & Schneider, 2015). Color and luminance can establish object correspondence, but it is unknown how these surface features contribute to transsaccadic visual processing. Here, we investigated whether changing the surface features color-and-luminance and color alone across saccades impairs postsaccadic object recognition. Participants made saccades to peripheral objects, which either maintained or changed their surface features across the saccade. After the saccade, participants briefly viewed a letter within the saccade target object (terminated by a pattern mask). Postsaccadic object recognition was assessed as participants' accuracy in reporting the letter. Experiment A used the colors green and red with different luminances as surface features, Experiment B blue and yellow with approximately the same luminances. Changing the surface features across the saccade deteriorated postsaccadic object recognition in both experiments. These findings reveal a link between object recognition and object correspondence relying on the surface features colors and luminance, which is currently not addressed in theories of transsaccadic perception. We interpret the findings within a recent theory ascribing this link to visual attention (Schneider, 2013).

Introduction

Human vision is based on a reiterating cycle of saccadic eye movements and intervals of relatively stable eye position, the so-called fixations. Saccades shift the eye rapidly, directing its foveal high-acuity region at potentially interesting parts of the environment. Fixations provide clear visual snapshots of objects, snapshots that are not corrupted by the suppression of information uptake or by the motion blur, which occur during saccades (e.g., Krock & Moore, 2014; Wurtz, 2008). However, snapshot-like sampling also poses a problem for perception and action. Perceiving objects as continuously present across saccades and to act based on this perception require that the snapshots of objects from successive fixations are linked (e.g., Higgins & Rayner, 2015; Schneider, 2013). This linkage is complicated by the fact that every saccade displaces an object's image on the retina and changes its resolution (e.g., Herwig & Schneider, 2014; Wurtz, 2008, 2015). How, then, does the visual system achieve coherent representations of external objects across saccades?

Current theories propose that coherent transsaccadic object representations depend on a test for object correspondence across saccades (Hollingworth, Richard, & Luck, 2008; Tas, Moore, & Hollingworth, 2012). This means the visual system tests whether postsaccadic and presaccadic object representations likely stem from the same external objects. If the test for object correspondence is positive (i.e., object correspondence is established), presaccadic object representations are updated with postsaccadic information (Demeyer, De Graef, Wagemans, & Verfaillie, 2009; Henderson & Anes, 1994), leaving only one postsaccadic representation (Tas et al., 2012; see also Schneider, 2013). Having

only one object representation may entail visual stability, the perception of a continuous and stable visual world despite the saccade-induced changes of retinal images (for reviews on visual stability, see Bridgeman, Van der Heijden, & Velichkovsky, 1994; Mathôt & Theeuwes, 2011; Wurtz, 2008; Wurtz, Joiner, & Berman, 2011). However, this also means that presaccadic and postsaccadic objects cannot be compared, which explains why object displacements (Bridgeman, Hendry, & Stark, 1975) and changes of visual object features (Deubel, Schneider, & Bridgeman, 2002; Weiß, Schneider, & Herwig, 2015) are hard to perceive when they occur during saccades. In contrast, if the test for object correspondence is negative (i.e., object correspondence is broken), presaccadic and postsaccadic objects are assumed to be represented separately (Deubel, Schneider, & Bridgeman, 1996; Schneider, 2013; Tas et al., 2012). This diminishes the perception of visual stability but helps to discriminate intrasaccadic object changes, presumably because the two representations can be compared (Deubel & Schneider, 1994; Deubel et al., 2002; Deubel et al., 1996; Tas et al., 2012; Weiß et al., 2015).

Transsaccadic object correspondence not only is important for visual stability but also has recently been shown to affect object recognition (Poth et al., 2015; see also Schneider, 2013). In this study, participants made saccades to a peripheral object. After the saccade, a letter was shown in this object and terminated by a pattern mask. Correspondence between the presaccadic and the postsaccadic object was broken with two different manipulations: first, by introducing a blank screen after eye landing and before onset of the postsaccadic object (see Deubel & Schneider, 1994; Deubel et al., 1996) and, second, by a large change of the luminance and the contrast polarity of the object during the saccade (see Tas et al., 2012). In both cases, recognition of the postsaccadic letter was deteriorated. This shows that breaking transsaccadic object correspondence impairs postsaccadic object recognition. Two explanations of this effect rely on the idea that the presaccadic and postsaccadic object are represented separately if object correspondence is broken. First, the theory of Task-dRiven visual Attention and working Memory (TRAM; Schneider, 2013) proposes that broken object correspondence (object continuity) across fixations results in two different object representations. Limited attentional resources must be split between the two representations, cutting the resources available to each individual object representation. This loss of attentional resources per object explains the deteriorated recognition of the postsaccadic object. Second, the creation of a separate postsaccadic representation in addition to the presaccadic one may delay (or hinder) processing of the postsaccadic object. Because the postsaccadic object was terminated by a mask, this delay would have become manifest in deteriorated recognition of the object (Poth et al., 2015).

Postsaccadic object recognition depends on transsaccadic object correspondence (Poth et al., 2015), but the mechanisms underlying this effect remain elusive. To shed light on these mechanisms, it is important to clarify which object features contribute to the test for object correspondence. Two classes of features are distinguished in the literature on object correspondence across occlusion (Hollingworth & Franconeri, 2009) and movement (Kahneman, Treisman, & Gibbs, 1992; Mitroff & Alvarez, 2007): spatiotemporal and surface features. Classical theories proposed that object correspondence was established solely (Kahneman et al., 1992) or primarily (Flombaum, Scholl, & Santos, 2009; Scholl, 2007) on the basis of spatiotemporal features. In stark contrast, however, more recent research revealed that object correspondence across occlusion can also be established based on surface features (such as color and/or luminance), even when it conflicts with the spatiotemporal feature location (Hollingworth & Franconeri, 2009). Along the same lines, object correspondence across saccades seems to rely on both spatiotemporal (Demeyer, De Graef, Wagemans, & Verfaillie, 2010; Deubel, Bridgeman, & Schneider, 1998; Deubel et al., 1996; Deubel et al., 2002) and surface features (Tas et al., 2012). As explained above, there is first evidence (Poth et al., 2015) that postsaccadic object recognition is deteriorated both when transsaccadic object correspondence is broken by blanking, which is a violation of spatiotemporal correspondence, and by introducing large changes of luminance and contrast polarity, which is a strong violation of surface feature correspondence. Critically, however, it remains to be clarified whether this holds also for surface features other than luminance and contrast polarity and less intense feature changes.

The surface feature of color is generally considered vital for human vision (e.g., Gegenfurtner & Kiper, 2003; Moutoussis, 2015), but it is unknown whether color is used for establishing object correspondence across saccades. Changing the apparent color of an object is a common manipulation to study how surface features contribute to object correspondence across occlusion (Hollingworth & Franconeri, 2009) and movement (Kahneman et al., 1992; Mitroff & Alvarez, 2007). However, such changes of apparent color may coincide with changes in luminance and contrast polarity (e.g., Mitroff & Alvarez, 2007), as these surface features are usually not distinguished from color. Therefore, the role of color for object correspondence across occlusion and movement remains unclear. The role of color may even be less clear for object correspondence across saccades. On the one hand, the color of a given object is represented with much lower

quality in the visual periphery than in the fovea (e.g., Hibino, 1992; Johnson, 1986; Livingstone & Hubel, 1987; Nagy & Wolf, 1993). Thus, if color was used to establish transsaccadic object correspondence, the natural differences between an object's peripheral presaccadic and its foveal postsaccadic color could erroneously break object correspondence. This would impair postsaccadic object recognition (Poth et al., 2015). One may therefore hypothesize that transsaccadic color changes are ignored. On the other hand, there is evidence that at least large changes in apparent color, which may include changes in luminance, can break transsaccadic object correspondence and perceived visual stability (Tas, 2015; cf. Hollingworth et al., 2008, for evidence from corrective saccades).

Here, we investigated whether breaking object correspondence across the saccade by changing the surface feature of color impairs postsaccadic object recognition. To retain the link to previous studies of object correspondence, we examined the effects of changes in apparent color (color, luminance, and contrast polarity) on transsaccadic object correspondence (Experiment A). In addition, we examined the effects of changes between approximately equiluminant colors on transsaccadic object correspondence (Experiment B). Both experiments employed the experimental paradigm by Poth et al. (2015; Experiment 2). Participants made saccades to a peripheral object, a letter was shown in this object after eye landing, and the letter presentation was terminated by a pattern mask. Participants' task was to report the identity of the postsaccadic letter. Transsaccadic object correspondence was manipulated in two conditions: The surface features of the object either stayed the same across the saccade (no-change condition) or they were changed during the saccade (change condition). Experiment A used the opponent colors green and red as surface features, each coinciding with a different physical luminance and contrast polarity. Experiment B used the colors blue and yellow with approximately the same luminance and contrast polarity. If breaking transsaccadic object correspondence by changing these surface features impairs postsaccadic object recognition, then letter report performance should be lower in the change compared with the no-change conditions of both experiments.

Method

Participants

Ten participants took part in Experiment A. They were between 20 and 30 years old ($MD = 25$ years), five were male, and five female. Ten different participants performed Experiment B. Their ages ranged from 21 to 26 years ($MD = 23.5$ years), and two were male and eight were female. All participants of both experiments reported normal color vision and normal or corrected-to-normal (contact lenses) visual acuity. All participants were paid and gave written informed consent before participation, and the experiments were approved by Bielefeld University's ethics committee.

Apparatus and stimuli

Participants performed the experiments in a dimly lit room. They viewed the 19-inch CRT screen (G90fB, ViewSonic, Brea, CA) from a distance of 71 cm while their head position was fixed by forehead and chin rests. The screen had a resolution of $1{,}024 \times 768$ pixels (at physical dimensions of 36×27 cm) and a refresh rate of 100 Hz, and it was controlled by a GeForce GT 640 (NVIDIA, Santa Clara, CA) graphics card. A video-based desktop-mounted eye tracker sampled participants' right eyes at 1000 Hz (Eyelink 1000, SR Research, Mississauga, Ontario, Canada). The eye tracker was calibrated using a nine-point grid procedure. Calibration was performed in the beginning of the experiment, after training trials, after a pause in about the middle of the experiment (and after participants had made 50 fixation or saccade errors in total). Saccades were detected online using velocity and acceleration thresholds of $35°$ (degrees of visual angle) $\times \, s^{-1}$ and $9500° \times s^{-2}$. Responses were collected using a standard QWERTZ computer keyboard. The experiment was programmed using the Psychophysics toolbox (3.0.12; Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997) and Eyelink toolbox (3.0.12; Cornelissen, Peters, & Palmer, 2002) extensions for MATLAB (R2014b; The MathWorks, Natick, MA).

Color and luminance were measured using an X-Rite i1 Pro spectrophotometer (Munich, Germany), and measurements are provided as CIE Lxy coordinates. A black ($L = 0.228$ cd/m^2, x = 0.290, y = 0.286) square ($0.1° \times 0.1°$) was used as central fixation stimulus. In Experiment A, saccade target objects were green ($L = 90.871$ cd/m^2, x = 0.279, y = 0.591) and red ($L = 30.664$ cd/m^2; x = 0.599, y = 0.327) ellipses ($0.65° \times 1.05°$). Note that the objects differed in luminance and contrast polarity. In Experiment B, saccade target objects had approximately the same luminance, and they were blue ($L = 37.113$ cd/m^2; x = 0.194; y = 0.143) and yellow ($L = 36.887$ cd/m^2; x = 0.415; y = 0.479). Letters (ABDEFGHJKLMNOPRSTVXZ; $0.32° \times 0.4°$) and special characters (%#§&; $0.4° \times 0.4°$) were written in Arial font and matched the gray background ($L = 47.687$ cd/m^2; x = 0.283, y = 0.291) in both experiments. In each experiment, 99 pattern masks were algorithmically produced for each individual
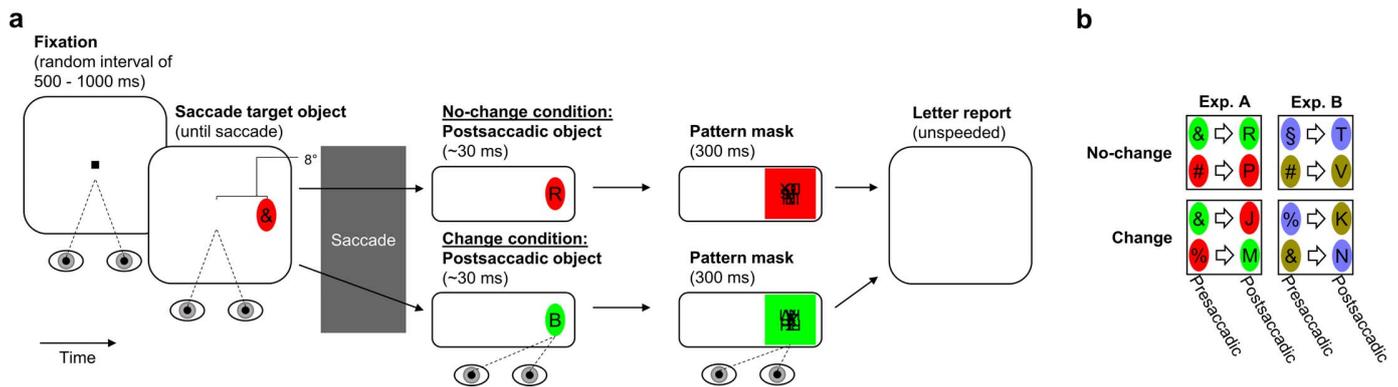
Figure 1. (a) Schematic illustration of the experimental paradigm. Participants fixated a central fixation stimulus, which was followed by an elliptic saccade target object, displaying one of two surface features (green and red with different luminances in Experiment A, blue and yellow with approximately the same luminance in Experiment B). This object appeared in the periphery and contained an irrelevant special character. Participants made a saccade to the object. In the no-change condition, the object displayed the same surface feature after the saccade. In the change condition, the postsaccadic object displayed different surface features than the presaccadic one. In both conditions, the postsaccadic object contained a letter and was terminated by a pattern mask. Participants reported the identity of the letter. Special characters, letters, and the background were gray (here drawn in black and white for better visibility). (b) Employed surface features and experimental conditions of Experiment A and Experiment B.

participant and for both colors. A large number of masks was used to minimize adaptation to them. The masks consisted of colored squares ($2° \times 2°$), filled with nine black letters that were drawn randomly without replacement from the set of letters. The nine letters were mirror reversed and upside down and overlapped partially, and all letters together covered an area of about $1° \times 1°$ within a square.

Procedure and design

Figure 1a illustrates the experimental paradigm, which is based on the paradigm by Poth et al. (2015). The participant pressed the space bar to start a trial. A central fixation stimulus was shown, and the participant fixated it for a random interval between 500 and 1000 ms. Next, an elliptic saccade target object appeared 8° horizontally from screen center. Whether the object appeared to the left or right of screen center was randomized across trials, whereby each side occurred equally often for each of the postsaccadic surface features and each condition. The object contained an irrelevant special character, which was randomly drawn from the set of special characters, and it was presented until the participant made a saccade to it. Figure 1b illustrates the surface features and experimental conditions of both experiments. In Experiment A, the presaccadic object was either green or red (coincident with different luminances and contrast polarities), and in Experiment B, it was either blue or yellow (with approximately the same luminances and contrast polarities). These surface features were randomized across trials, each occurring equally

often in each condition. The postsaccadic object contained a letter that was randomly drawn from the set of letters[1] and was shown during the saccade (on the next screen refresh after detection of saccade onset). The presaccadic and the postsaccadic object had the same surface features in the no-change condition and different surface features in the change condition. Thus, in the change condition of Experiment A, green objects changed into red ones and vice versa. In the change condition of Experiment B, blue objects changed into yellow ones and vice versa. After the postsaccadic object, a pattern mask of the same surface feature was presented. This mask was shown two or three screen refreshes (frames) after the online detection of saccade end (and its registration by the experimental software), so that the postsaccadic object was visible after the saccade end detection for 31 ms on average ($SD = 3$ ms). The mask was randomly drawn from the set of produced masks and was shown for 300 ms. It was followed by a blank screen, and participants reported the identity of the letter using the keyboard. There was no time limit for the report. The next trial could be started after an intertrial interval of 100 ms. Participants did not receive any instructions regarding the surface features or changes.

Participants performed 64 trials of each condition in randomized order. Trials were aborted and repeated on a randomly selected subsequent trial if observers did not fixate the central fixation cross or missed the saccade target object by more than 2.5°. In Experiment A, 24.3% of the trials were repeated, in Experiment B 31.11%. Participants performed 32 training trials before each experiment.

	Letter report performance (proportion correct)	Saccade latency (ms)	Saccade landing error (distance in °)
Experiment A			
No-change, green	0.96 (0.08)	149 (12)	0.80 (0.13)
No-change, red	0.83 (0.20)	153 (12)	0.74 (0.13)
Change, green	0.86 (0.16)	153 (12)	0.74 (0.16)
Change, red	0.60 (0.27)	148 (10)	0.76 (0.19)
Experiment B			
No-change, blue	0.80 (0.25)	168 (19)	0.93 (0.29)
No-change, yellow	0.72 (0.27)	159 (16)	0.90 (0.25)
Change, blue	0.71 (0.26)	159 (14)	0.88 (0.28)
Change, yellow	0.63 (0.31)	167 (16)	0.91 (0.27)

Table 1. Means of letter report performance, saccade latency, and saccade landing errors across participants for Experiment A and B. Values are provided for cells formed by the two conditions (no-change and change) and the two postsaccadic surface features (green and red in Experiment A, and blue and yellow in Experiment B). Standard deviations are in parentheses.

Results and discussion

Trials were excluded from analysis if saccade latency (the time from onset of the saccade target object until saccade onset detection) was less than 100 ms (anticipatory saccades) or greater than 400 ms. Two trials were excluded from Experiment A and four trials from Experiment B. Table 1 provides descriptive statistics of all dependent variables in both experiments.

Letter report performance in Experiment A

Letter report performance was assessed as the proportion of correctly reported letters for each individual participant. Because of the truncated range proportions take, all analyses were also performed on acrsine-squareroot-transformed proportions in addition to original proportions. Both sets of analyses yielded consistent results, and therefore, only analyses of original proportions are reported.

Figure 2a depicts the mean proportion of correctly reported letters across participants of Experiment A. The effects of condition and postsaccadic surface features on letter report performance were analyzed using a 2×2 (no-change vs. change $\times$ green vs. red) repeated-measures analysis of variance (ANOVA; with type III sums of squares and η_G^2 as effect size; Bakeman, 2005). The ANOVA revealed a significant main effect of condition, $F(1, 9) = 15.607$, $p = 0.003$, $\eta_G^2 = 0.181$. Performance was higher in the no-change ($M = 0.90$, $SD = 0.13$) compared with the change condition ($M = 0.73$, $SD = 0.19$). Thus, breaking transsaccadic object correspondence by changing the combination of the surface features color, luminance, and contrast polarity impaired recognition of the postsaccadic letter. As argued previously (Poth et al., 2015), one might suspect that the change of the special character in the presaccadic object into the letter in the postsaccadic object also broke transsaccadic object correspondence. It is important, however, that even if this were the case, the present results would still demonstrate an effect of breaking object correspondence by changing color, luminance, and contrast polarity in addition to the possible effect of changing the presaccadic special character.

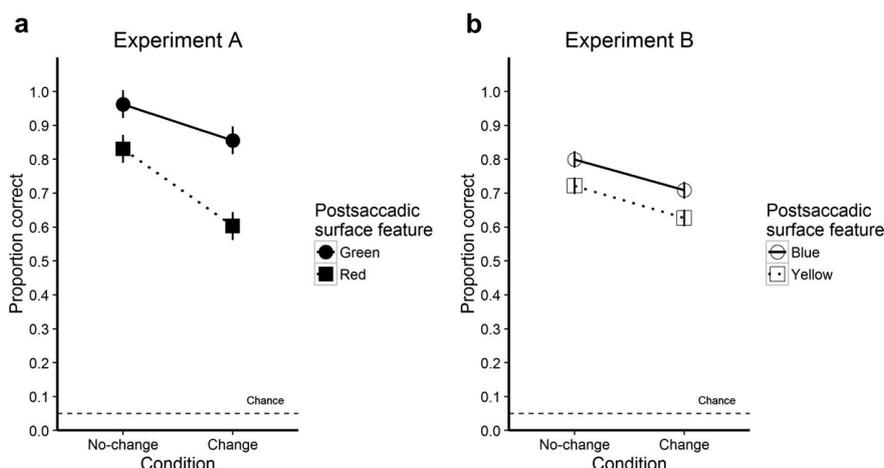Figure 2. Letter report performance. Depicted are mean proportions of correct letter reports in the two conditions (change vs. no change) and for both postsaccadic surface features (green and red in Experiment A, blue and yellow in Experiment B). Error-bars indicate ±1 standard error of the mean for within-subjects designs (Loftus & Masson, 1994); the dashed line indicates chance level. (a) Experiment A. (b) Experiment B.

There was also a significant main effect of post-saccadic surface features $F(1, 9) = 12.751$, $p = 0.006$, $\eta_G^2 = 0.227$, showing that performance was higher for green ($M = 0.91$, $SD = 10.10$) than for red ($M = 0.72$, $SD = 0.22$) postsaccadic objects. Postsaccadic object color, luminance, and contrast polarity may have affected the visibility and hence recognition of the postsaccadic letter (as has been shown for luminance contrast by Petersen & Andersen, 2012). The interaction between the two factors was significant as well, $F(1, 9) = 9.895$, $p = 0.012$, $\eta_G^2 = 0.028$. Specifically, the difference between performance in the no-change compared with the change condition was smaller for the green postsaccadic objects ($M = 0.11$, $SD = 0.14$) compared with the red ones ($M = 0.23$, $SD = 0.16$).

Saccade latencies and landing errors in Experiment A

Saccade latencies were assessed as each participant's mean interval (in ms) between the onset of the saccade target object and detection of the saccade. Table 1 provides descriptive statistics for saccade latencies in the two conditions in conjunction with the postsaccadic surface features. Saccade latency was affected neither by condition, $F(1, 9) = 0.067$, $p = 0.801$, $\eta_G^2 < 0.001$, nor by postsaccadic surface features, $F(1, 9) = 0.214$, $p = 0.655$, $\eta_G^2 < 0.001$. However, there was a disordinal interaction between the two factors, $F(1, 9) = 21.152$, $p = 0.001$, $\eta_G^2 = 0.036$. Note that this interaction effect corresponds to a main effect of presaccadic surface features if the presaccadic rather than the postsaccadic surface features entered the ANOVA as second factor besides condition. Saccade latencies were shorter when presaccadic objects were green ($M = 148$ ms; $SD = 11$ ms) rather than red ($M = 153$ ms; $SD = 12$ ms). This may indicate that green saccade target objects were perceptually more salient, which implies they were easier to detect and localize than red ones, leading to faster saccades for the former compared with the latter.

Saccade landing errors were assessed as each participant's mean Euclidian distance (in °) between saccade landing sites and saccade target objects (i.e., the center coordinates of these objects). Neither condition, $F(1, 9) = 0.362$, $p = 0.562$, $\eta_G^2 = 0.003$, nor postsaccadic surface features, $F(1, 9) = 1.399$, $p = 0.267$, $\eta_G^2 = 0.007$, nor the interaction between the two factors, $F(1, 9) = 2.346$, $p = 0.160$, $\eta_G^2 = 0.015$, had significant effects on saccade landing errors.

Letter report performance in Experiment B

As for Experiment A, letter report performance was analyzed based on the original and the arcsine-square-root-transformed proportions of correct reports. Only the analyses of the original proportions are reported

because both sets of analyses delivered consistent results.

Figure 2b depicts the mean proportion of correctly reported letters across participants of Experiment B. The ANOVA showed a significant main effect of condition, $F(1, 9) = 13.514$, $p = 0.005$, $\eta_G^2 = 0.031$. Thereby, performance was higher in the no-change ($M = 0.76$; $SD = 0.26$) than in the change condition ($M = 0.67$; $SD = 0.28$). This indicates that breaking trans-saccadic object correspondence by changing between the approximately equiluminant object colors impaired recognition of the postsaccadic letter.

There was also a significant main effect of post-saccadic color, $F(1, 9) = 7.902$, $p = 0.020$, $\eta_G^2 = 0.023$, whereby letter report performance was higher for blue ($M = 0.75$; $SD = 0.25$) than for yellow postsaccadic objects ($M = 0.67$; $SD = 0.29$). This may suggest that the postsaccadic color affected the visibility and recognition of the letter. The interaction between condition and postsaccadic color was not significant, $F(1, 9) = 0.013$, $p = 0.912$, $\eta_G^2 < 0.001$.

Saccade latencies and landing errors in Experiment B

Saccade latency was unaffected by condition, $F(1, 9) = 0.079$, $p = 0.785$, $\eta_G^2 < 0.001$, and postsaccadic color, $F(1, 9) = 0.026$, $p = 0.876$, $\eta_G^2 < 0.001$. However, there was an interaction between these two factors, $F(1, 9) = 45.035$, $p < 0.001$, $\eta_G^2 = 0.067$. This interaction effect corresponds to a main effect of presaccadic color if this was included in the ANOVA instead of the postsaccadic color. Saccades were faster when presaccadic objects were yellow ($M = 159$ ms; $SD = 15$ ms) rather than blue ($M = 167$ ms; $SD = 17$ ms). This effect may be due to a higher perceptual saliency of the yellow compared with the blue objects, which may have sped up the detection and localization of saccade target objects (see Experiment A).

Saccade landing errors were neither affected by condition, $F(1, 9) = 0.377$, $p = 0.554$, $\eta_G^2 = 0.001$, nor by postsaccadic color, $F(1, 9) < 0.001$, $p = 0.987$, $\eta_G^2 < 0.001$, nor by the interaction of the two factors $F(1, 9) = 0.949$, $p = 0.355$, $\eta_G^2 = 0.003$.

General discussion

We tested the hypothesis that breaking object correspondence across the saccade by changing surface features impairs postsaccadic object recognition. The results of two experiments support this hypothesis. Experiment A revealed that transsaccadic changes between the combined surface features color, luminance, and contrast polarity deteriorate postsaccadic object recognition. Experiment B yielded similar

findings for colors of approximately the same physical luminance and the same contrast polarity. Together, the results indicate that postsaccadic object recognition depends on mechanisms establishing transsaccadic object correspondence on the basis of these surface features. Although it has been shown before that luminance and contrast polarity contribute to transsaccadic object correspondence, the effects of color with approximate equiluminance may be surprising. That is, the presaccadic peripheral and the postsaccadic foveal retinal images of an object provide color information of substantially different quality (e.g., Hibino, 1992; Johnson, 1986; Nagy & Wolf, 1993; and possibly perceived luminance, Livingstone & Hubel, 1987). Therefore, color might not be an ideal feature for establishing object correspondence across the saccade. Nevertheless, color seems to be used for this purpose in concert with luminance and contrast polarity, together paving the way for object recognition.

Hitherto, the dependency of postsaccadic object recognition on transsaccadic object correspondence has only been studied using two correspondence-breaking manipulations: blanking and changing the contrast polarity of achromatic objects, which coincided with large luminance changes (Poth et al., 2015). The present findings extend these results to chromatic objects. In Experiment A, transsaccadic object correspondence was broken by changing between colors with different physical luminances and contrast polarities. Experiment B replicated the results using colored objects of about the same physical luminance. The perceived luminance of a given color differs at different retinal locations (and this might depend on individual participants; Livingstone & Hubel, 1987), so that saccades might always imply a change of an object's perceived luminance. Consequently, transsaccadic changes of object color may change perceived luminance, even for physically equiluminant objects. This means that perceived luminance could still have contributed to the effect of Experiment B. Hence, this effect may either be due to changes of the object's chromaticity and/or the associated luminance changes. In either way, this demonstrates that the changes are not ignored when transsaccadic object correspondence is determined, so that they affect postsaccadic object recognition. The present findings indicate that postsaccadic object recognition depends on mechanisms of transsaccadic object correspondence that use information from surface features in general or at least from the surface features of contrast polarity and luminance and of color, whereby the latter might inherently include a contribution of luminance. As such, these findings conflict with the view that transsaccadic object correspondence relies exclusively on spatiotemporal features (Kahneman et al., 1992; with respect to effects

on postsaccadic object recognition). In sum, our experiments demonstrated that postsaccadic object recognition is deteriorated when transsaccadic object correspondence is broken by changes of spatiotemporal features (blanking) and of surface features such as contrast polarity and luminance (Poth et al., 2015), combined color, luminance, and contrast polarity (Experiment A) and of color (Experiment B).

Presaccadic and postsaccadic objects are assumed to be represented as separate entities if transsaccadic object correspondence is broken (e.g., Tas et al., 2012; see also Deubel et al., 1996; Schneider, 2013). Consequently, the deteriorated object recognition can be interpreted in at least two ways, which need not be mutually exclusive. First, creating a separate representation for the postsaccadic object may delay or hinder processing of this object. Recognition of this object would then be deteriorated, especially if the object is terminated by a mask (Poth et al., 2015). Second, having separate representations of the presaccadic and the postsaccadic object may introduce attentional competition between them (Schneider, 2013). That is, limited attentional resources are split between the two representations. Fewer resources are available for processing each individual object representation, which then deteriorates object recognition (Bundesen, 1990; Desimone & Duncan, 1995). We assessed postsaccadic object recognition as performance in reporting a letter, which participants viewed after the saccade in the saccade target object. The letter should have been processed as part of the surface of this object, so that letter report performance reflects recognition of this object (cf. Henderson & Anes, 1994; Kahneman et al., 1992; Mitroff & Alvarez, 2007). However, it might be possible that the letter has been processed as a separate object. Letter report performance would then reflect recognition of a newly appearing object at the spatial location of the saccade target object rather than recognition of this object itself. This would still be in line with the two explanations, following the assumption of competitive object recognition (e.g., Bundesen, Habekost, & Kyllingsbæk, 2005; Desimone & Duncan, 1995). Either the creation of a representation for the letter at this very location would be delayed or it would suffer from greater attentional competition if transsaccadic object correspondence was broken and led to two rather than one representation of the saccade target object.

The present findings argue that the surface features of combined color, luminance, and contrast polarity, as well as color alone, are used by object correspondence mechanisms, which track objects across saccade-induced shifts of retinal images. Moreover, they argue that these object correspondence mechanisms affect mechanisms of object recognition. Based on TRAM (Schneider, 2013), we suggest that this may be due to an

interface between the two sets of mechanisms, which is provided by visual attention (see also Poth et al., 2015). We elaborate this hypothesis in the following.

Which mechanism tracks objects across the saccade-induced changes of retinal images? The tracking of object locations across saccades may be accomplished by retinotopically organized brain areas (as, e.g., monkeys' lateral intraparietal area, Duhamel, Colby, & Goldberg, 1992; superior colliculus, Walker, Fitzgibbon, & Goldberg, 1995; and frontal eye fields, Umeno & Goldberg, 1997). Shortly before a saccade, neurons in these areas respond to stimuli at the locations where their receptive fields (the retinal regions from which they receive information) will be after the saccade. This has been interpreted as a *predictive remapping* of the neurons' receptive fields to these locations (Duhamel et al., 1992; but see Zirnsak, Steinmetz, Noudoost, Xu, & Moore, 2014). The necessary information about the amplitude and direction of the saccade seems to come from a corollary discharge (efference copy) of the motor signals eliciting the saccade (Sommer & Wurtz, 2006). When the receptive field of a neuron is predictively remapped, the neuron responds to a particular object before the saccade. The following saccade-induced shift of the neuron's receptive field makes the neuron respond to the same object again after the saccade. An additional process comparing the presaccadic and postsaccadic activity of such neurons might then allow one to infer the presence of an object before and after the saccade. Therefore, such a comparison has been hypothesized to underlie the perception of visual stability of object locations across saccades (Cavanaugh, Berman, Joiner, & Wurtz, 2016; Duhamel et al., 1992; Wurtz et al., 2011). The comparison may be part of the neuronal implementation of the test for transsaccadic object correspondence, the test that is assumed to govern visual stability (Poth et al., 2015; Tas et al., 2012). One problem remains, however. The comparison provides information about whether an object is present at a given location before and after the saccade. It does not provide information about the (surface) features of the object (e.g., Cavanagh, Hunt, Afraz, & Rolfs, 2010). Surface features clearly contribute to visual stability (Tas, 2015; Tas et al., 2012), which argues that the test for transsaccadic object correspondence cannot be accomplished based on the described comparison alone. A potential solution to this problem is provided by TRAM (Schneider, 2013).

TRAM proposes a mechanism that tests for object correspondence (object continuity) across interruptions of visual input in between fixations (as due to the suppression of input during saccades, e.g., Krock & Moore, 2014) and across changes of visual objects within a fixation. Critically, this test for object correspondence allows us to take into account the (surface) features of objects. Correspondence between objects of successive fixations should be tested on the basis of the objects' attentional weights (Schneider, 2013). The concept of attentional weight comes from Bundesen's (1990) theory of visual attention. The attentional weight of an object indicates its current relevance in a spatially organized fashion. It is computed as the sum of the sensory evidences that the object has certain features, whereby the sensory evidence for each feature is weighted by the current relevance of this feature. According to TRAM, the attentional weight that an object will have after the saccade should be predicted before the saccade is executed. This counteracts the changes of attentional weights due to predictable changes of sensory evidence, for example, due to saccade-induced shifts of retinal images of objects, which change their resolution (cf. Herwig & Schneider, 2014). After the saccade, the predicted attentional weight is compared with the current attentional weight of the object. The test for object correspondence is positive if the two agree and negative if they disagree. Evidence that prediction affects attentional weights has been provided recently (Poth, Petersen, Bundesen, & Schneider, 2014). In this study, participants increased the attentional weight of an object that was monitored for a luminance change in order to compensate for a low predicted (expected) salience of this change.

Attentional weights are proposed to be implemented in priority maps (Bundesen et al., 2005): spatially organized (retinotopic) brain areas whose neurons seem to code for the relevance and physical salience of objects (Bisley & Goldberg, 2010; Fecteau & Munoz, 2006; Zelinsky & Bisley, 2015). Interestingly, the brain areas supposed to contain priority maps are among the ones whose neurons seem to predictively remap their receptive fields before saccades (e.g., monkeys' lateral intraparietal area, Duhamel et al., 1992; superior colliculus, Walker et al., 1995; frontal eye fields, Umeno & Goldberg, 1997; note that some extrastriate areas show remapping as well, Nakamura & Colby, 2002). Extending TRAM (Schneider, 2013), we therefore hypothesize that predictive remapping contributes to the prediction of attentional weights.

Now that we have sketched a mechanism using attentional weights to test for transsaccadic object correspondence, we can ask how transsaccadic object correspondence is linked to the object recognition system. An answer to this question may reside in the attentional weights as well. In fact, attentional weights have originally been introduced to explain how relevant objects are selected for object recognition at the expense of irrelevant ones (Bundesen, 1990). To be recognized, an object has to be processed with enough processing resources, which might mean by enough neurons (Bundesen et al., 2005; to eventually enter

visual working memory where recognition is complete and report possible). Attentional weights control the allocation of processing resources to objects; each object is assumed to receive processing resources in proportion to its attentional weight relative to the summed attentional weights of all objects in the visual field. More neurons are allocated to relevant than irrelevant objects by virtue of a gating mechanism: Gates in between the lower and higher cortical levels of the ventral object recognition pathway are opened and closed so that the receptive fields of neurons are *dynamically remapped* to locations of relevant objects (Bundesen et al., 2005; Desimone & Duncan, 1995). Importantly, which gates are opened and which are closed is determined by the attentional weights from spatially organized priority maps (Bundesen et al., 2005).

Taken together, following TRAM (Schneider, 2013; see also Poth et al., 2015), transsaccadic object correspondence and object recognition should be linked by attentional weights. In this view, postsaccadic object recognition may be impaired because of broken transsaccadic object correspondence for (at least) two reasons. First, a negative test for object correspondence means that an object's predicted attentional weight and its actual postsaccadic attentional weight mismatch. Therefore, after the saccade, there are two discrepant attentional weights present. Postsaccadic object recognition may be impaired because the postsaccadic object is allocated neuronal processing resources according to its attentional weight divided by the sum of the two present attentional weights. In contrast, if the test for object correspondence is positive, this means that the object's predicted and postsaccadic attentional weight match. In this case, there is only one postsaccadic attentional weight. Hence, the postsaccadic object is allocated all available neuronal processing resources (i.e., according to its attentional weight divided by only itself). Compared with the situation of a negative test for object correspondence, the object receives more processing resources, which consequently improves object recognition. In addition, the state of the gates in the ventral object recognition pathway that is dictated by the predicted postsaccadic attentional weight is then the same as the one dictated by the actual postsaccadic attentional weight. This may provide a basis for transsaccadic updating processes (cf. Demeyer et al., 2009; Henderson & Anes, 1994; and transsaccadic integration, Ganmor, Landy, & Simoncelli, 2015; Herwig, 2015; Wolf & Schütz, 2015; but see Wittenberg, Bremmer, & Wachtler, 2008). The lower- or mid-level (surface) features of the postsaccadic object may be routed to presaccadically created object representations through a consistent state of gates (cf. Poth et al., 2015). As a consequence, representations from the presaccadic and postsaccadic retinal images of objects

are combined within a common postsaccadic object representation. This leads to the second reason why breaking transsaccadic object correspondence affects postsaccadic object recognition. We suggest that combining presaccadic and postsaccadic representations may provide computational savings: The processing of the object in question that started before the saccade can be continued after the saccade. In contrast, if object correspondence is broken and updating is blocked, then processing of the postsaccadic object might have to start completely anew. As a consequence, object recognition might be delayed, leading to performance decrements (especially when postsaccadic objects are terminated by masks; Poth et al., 2015).

In sum, extending TRAM (Schneider, 2013), we hypothesize that the dynamic remapping of receptive fields in the ventral stream for object recognition (for reviews, see Bundesen et al., 2005; Desimone & Duncan, 1995) and the predictive remapping of receptive fields in dorsal and frontal areas (for a review, see Wurtz et al., 2011) is accomplished by common attentional weights. In this vein, attentional weights allow us to keep track of objects across saccades, they establish correspondence between presaccadic and postsaccadic objects, and they distribute neuronal resources across these objects for object recognition.

Conclusion

The present study investigated whether breaking object correspondence across the saccade by changing the surface features of combined color, luminance, and contrast polarity and the surface feature of color impair postsaccadic object recognition. The findings from two experiments indicate that this is the case. As such, they provide new evidence for an interface between mechanisms of transsaccadic object correspondence relying on surface features and mechanisms of object recognition. Based on the TRAM theory (Schneider, 2013), we propose that this interface is provided by visual attention.

Keywords: saccadic eye movements, visual attention, visual stability, object recognition, transsaccadic memory

Acknowledgments

Commercial relationships: none.
Corresponding author: Christian H. Poth.
Email: c.poth@uni-bielefeld.de.
Address: Neuro-cognitive Psychology, Department of Psychology, Bielefeld University, Bielefeld, Germany.

Footnote

[1] Letters and special characters were drawn randomly with a unique random sequence per participant. χ^2 tests indicated that the presentation frequencies of specific letters (Experiment A: $\chi^2[57] = 65.578$, $p = 0.204$; Experiment B: $\chi^2[57] = 56.006$, $p = 0.512$) and of the combination of special characters and letters (Experiment A: $\chi^2[237] = 238.86$, $p = 0.454$; Experiment B: $\chi^2[237] = 215.84$, $p = 0.834$) did not significantly depend on the cells formed by the two experimental conditions (no-change and change) and the two postsaccadic surface features (green and red in Experiment A and blue and yellow in Experiment B).

References

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*, 379–384, doi:10.3758/BF03192707.

Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience*, *33*, 1–21, doi:10.1146/annurev-neuro-060909-152823.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436, doi:10.1163/156856897X00357.

Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research*, *15*, 719–722, doi:10.1016/0042-6989(75)90290-4.

Bridgeman, B., Van der Heijden, A. H. C., & Velichkovsky, B. M. (1994). A theory of visual stability across saccadic eye movements. *Behavioral and Brain Sciences*, *17*, 247–292, doi:10.1017/S0140525X00034361.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*, 523–547, doi:10.1037/0033-295X.97.4.523.

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review*, *112*, 291–328, doi:10.1037/0033-295X.112.2.291.

Cavanagh, P., Hunt, A. R., Afraz, A., & Rolfs, M. (2010). Visual stability based on remapping of attention pointers. *Trends in Cognitive Sciences*, *14*, 147–153, doi:10.1016/j.tics.2010.01.007.

Cavanaugh, J., Berman, R. A., Joiner, W. M., & Wurtz, R. H. (2016). Saccadic corollary discharge underlies stable visual perception. *Journal of Neuroscience*, *36*, 31–42, doi:10.1523/JNEUROSCI.2054-15.2016.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*, 613–617, doi:10.3758/BF03195489.

Demeyer, M., & De Graef, P., Wagemans, J., & Verfaillie, K. (2009). Transsaccadic identification of highly similar artificial shapes. *Journal of Vision*, *9*(4):28, 1–14, doi:10.1167/9.4.28. [PubMed] [Article]

Demeyer, M., & De Graef, P., Wagemans, J., & Verfaillie, K. (2010). Object form discontinuity facilitates displacement discrimination across saccades. *Journal of Vision*, *10*(6):17, 1–14, doi:10.1167/10.6.17. [PubMed] [Article]

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222, doi:10.1146/annurev.ne.18.030195.001205.

Deubel, H., Bridgeman, B., & Schneider, W. X. (1998). Immediate post-saccadic information mediates space constancy. *Vision Research*, *38*, 3147–3159, doi:10.1016/S0042-6989(98)00048-0.

Deubel, H., & Schneider, W. X. (1994). Perceptual stability and postsaccadic visual information: Can man bridge a gap? *Behavioral and Brain Sciences*, *17*, 259, doi:10.1017/S0140525X00034397.

Deubel, H., Schneider, W. X., & Bridgeman, B. (1996). Postsaccadic target blanking prevents saccadic suppression of image displacement. *Vision Research*, *36*, 985–996, doi:10.1016/0042-6989(95)00203-0.

Deubel, H., Schneider, W. X., & Bridgeman, B. (2002). Transsaccadic memory of position and form. *Progress in Brain Research*, *140*, 165–180, doi:10.1016/S0079-6123(02)40049-0.

Duhamel, J.-R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, *255*, 90–92, doi:10.1126/science.1553535.

Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: A priority map for target selection. *Trends in Cognitive Sciences, 10*, 382–390, doi:10.1016/j.tics.2006.06.011.

Flombaum, J. I., Scholl, B. J., & Santos, L. R. (2009). Spatiotemporal priority as a fundamental principle of object persistence. In B. Hood & L. Santos (Eds.), *The origins of object knowledge* (pp. 135–164). Oxford, UK: Oxford University Press.

Ganmor, E., Landy, M. S., & Simoncelli, E. P. (2015). Near-optimal integration of orientation information across saccades. *Journal of Vision, 15*(16):8, 1–12, doi:10.1167/15.16.8. [PubMed] [Article]

Gegenfurtner, K. R., & Kiper, D. C. (2003). Color vision. *Annual Review of Neuroscience, 26*, 181–206, doi:10.1146/annurev.neuro.26.041002.131116.

Henderson, J. M., & Anes, M. D. (1994). Roles of object-file review and type priming in visual identification within and across eye fixations. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 826–839, doi:10.1037/0096-1523.20.4.826.

Herwig, A. (2015). Transsaccadic integration and perceptual continuity. *Journal of Vision, 15*(16):7, 1–6, doi:10.1167/15.16.7. [PubMed] [Article]

Herwig, A., & Schneider, W. X. (2014). Predicting object features across saccades: Evidence from object recognition and visual search. *Journal of Experimental Psychology: General, 143*, 1903–1922, doi:10.1037/a0036781.

Hibino, H. (1992). Red-green and yellow-blue opponent-color responses as a function of retinal eccentricity. *Vision Research, 32*, 1955–1964, doi:10.1016/0042-6989(92)90055-N.

Higgins, E., & Rayner, K. (2015). Transsaccadic processing: Stability, integration, and the potential role of remapping. *Attention, Perception & Psychophysics, 77*, 3–27, doi:10.3758/s13414-014-0751-y.

Hollingworth, A., & Franconeri, S. L. (2009). Object correspondence across brief occlusion is established on the basis of both spatiotemporal and surface feature cues. *Cognition, 113*, 150–166, doi:10.1016/j.cognition.2009.08.004.

Hollingworth, A., Richard, A. M., & Luck, S. J. (2008). Understanding the function of visual short-term memory: Transsaccadic memory, object correspondence, and gaze correction. *Journal of Experimental Psychology: General, 137*, 163–181, doi:10.1037/0096-3445.137.1.163.

Johnson, M. A. (1986). Color vision in the peripheral retina. *American Journal of Optometry & Physiological Optics, 63*, 97–103.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 175–219, doi:10.1016/0010-0285(92)90007-O.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception, 36*(Abstract Supplement).

Krock, R. M., & Moore, T. (2014). The influence of gaze control on visual perception: Eye movements and visual stability. *Cold Spring Harbor Symposia on Quantitative Biology, 79*, 123–130, doi:10.1101/sqb.2014.79.024836.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience, 7*, 3416–3468.

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*, 476–490, doi:10.3758/BF03210951.

Mathôt, S., & Theeuwes, J. (2011). Visual attention and stability. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*, 516–527, doi:10.1098/rstb.2010.0187.

Mitroff, S. R., & Alvarez, G. A. (2007). Space and time, not surface features, guide object persistence. *Psychonomic Bulletin & Review, 14*, 1199–1204, doi:10.3758/BF03193113.

Moutoussis, K. (2015). The physiology and psychophysics of the color-form relationship: A review. *Frontiers in Psychology, 6*, 1407, doi:10.3389/fpsyg.2015.01407.

Nagy, A. L., & Wolf, S. (1993). Red-green color discrimination in peripheral vision. *Vision Research, 33*, 235–242, doi:10.1016/0042-6989(93)90161-O.

Nakamura, K., & Colby, C. L. (2002). Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proceedings of the National Academy of Sciences, USA, 99*, 4026–4031, doi:10.1073/pnas.052379899.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442, doi:10.1163/156856897X00366.

Petersen, A., & Andersen, T. S. (2012). The effect of exposure duration on visual character identification in single, whole, and partial report. *Journal of Experimental Psychology: Human Perception and Performance, 38*, 498–514, doi:10.1037/a0026728.

Poth, C. H., Herwig, A., & Schneider, W. X. (2015). Breaking object correspondence across saccadic eye

movements deteriorates object recognition. *Frontiers in Systems Neuroscience*, *9*, 176, doi:10.3389/fnsys.2015.00176.

Poth, C. H., Petersen, A., Bundesen, C., & Schneider, W. X. (2014). Effects of monitoring for visual events on distinct components of attention. *Frontiers in Psychology*, *5*, 930, doi:10.3389/fpsyg.2014.00930.

Schneider, W. X. (2013). Selective visual processing across competition episodes: A theory of task-driven visual attention and working memory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*, 20130060, doi:10.1098/rstb.2013.0060.

Scholl, B. J. (2007). Object persistence in philosophy and psychology. *Mind & Language*, *22*, 563–591, doi:10.1111/j.1468-0017.2007.00321.x.

Sommer, M. A., & Wurtz, R. H. (2006). Influence of the thalamus on spatial visual processing in frontal cortex. *Nature*, *444*, 374–377, doi:10.1038/nature05279.

Tas, A. C. (2015). *The role of visual stability in representations of pre- and post-saccadic objects* (Unpublished doctoral thesis). University of Iowa. Retrieved from http://ir.uiowa.edu/etd/1916

Tas, A. C., Moore, C. M., & Hollingworth, A. (2012). An object-mediated updating account of insensitivity to transsaccadic change. *Journal of Vision*, *12*(11):18, 1–13, doi:10.1167/12.11.18. [PubMed] [Article]

Umeno, M. M., & Goldberg, M. E. (1997). Spatial processing in the monkey frontal eye field. I. Predictive visual responses. *Journal of Neurophysiology*, *78*, 1373–1383.

Walker, M. F., Fitzgibbon, E. J., & Goldberg, M. E. (1995). Neurons in the monkey superior colliculus predict the visual result of impending saccadic eye movements. *Journal of Neurophysiology*, *73*, 1988–2003.

Weiß, K., Schneider, W. X., & Herwig, A. (2015). A "blanking effect" for surface features: Transsaccadic spatial-frequency discrimination is improved by postsaccadic blanking. *Attention, Perception & Psychophysics*, *77*, 1500–1506, doi:10.3758/s13414-015-0926-1.

Wittenberg, M., Bremmer, F., & Wachtler, T. (2008). Perceptual evidence for saccadic updating of color stimuli. *Journal of Vision*, *8*(14):9, 1–9, doi:10.1167/8.14.9. [PubMed] [Article]

Wolf, C., & Schütz, A. C. (2015). Trans-saccadic integration of peripheral and foveal feature information is close to optimal. *Journal of Vision*, *15*(16):1, 1–18, doi:10.1167/15.16.1. [PubMed] [Article]

Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Research*, *48*, 2070–2089, doi:10.1016/j.visres.2008.03.021.

Wurtz, R. H. (2015). Brain mechanisms for active vision. *Daedalus*, *144*, 10–21, doi:10.1162/DAED_a_00314.

Wurtz, R. H., Joiner, W. M., & Berman, R. A. (2011). Neuronal mechanisms for visual stability: Progress and problems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 492–503, doi:10.1098/rstb.2010.0186.

Zelinsky, G. J., & Bisley, J. W. (2015). The what, where, and why of priority maps and their interactions with visual working memory. *Annals of the New York Academy of Sciences*, *1339*, 154–164, doi:10.1111/nyas.12606.

Zirnsak, M., Steinmetz, N. A., Noudoost, B., Xu, K. Z., & Moore, T. (2014). Visual space is compressed in prefrontal cortex before eye movements. *Nature*, *507*, 504–507, doi:10.1038/nature13149.

Study 3

Attentional competition across saccadic eye movements
(submitted to the Journal of Experimental Psychology:
Human Perception and Performance)

1

2

3

4   Attentional competition across saccadic eye movements

5   Christian H. Poth & Werner X. Schneider

6   Bielefeld University

7

8

9   Author note

10   Christian H. Poth, Neuro-cognitive Psychology, Department of Psychology, and

11 Cluster of Excellence Cognitive Interaction Technology, Bielefeld University. P.O. box

12 100131, 33501 Bielefeld, Germany.

13   Werner X. Schneider, Neuro-cognitive Psychology, Department of Psychology, and

14 Cluster of Excellence Cognitive Interaction Technology, Bielefeld University. P.O. box

15 100131, 33501 Bielefeld, Germany.

16   Correspondence concerning this article should be addressed to Christian H. Poth,

17 Neuro-cognitive Psychology, Department of Psychology, and Cluster of Excellence Cognitive

18 Interaction Technology, Bielefeld University. P.O. box 100131, 33501 Bielefeld, Germany.

19 E-mail: c.poth@uni-bielefeld.de

26   Words: 5539, figures: 3, tables: 1

Abstract

Human behavior is guided by visual object recognition. For being recognized, objects

compete for limited attentional processing resources. The more objects compete, the lower is

each objects' processing speed. Here, we ask whether this competition is confined to eye

fixations, periods of relatively stable gaze, or whether it extends from one fixation to the next,

across the saccadic eye movements. Participants made saccades to a peripheral saccade target.

They reported a letter that was shown after the saccade within the saccade target and for

different durations (mask-terminated). Processing speed of this letter was measured by

modeling report performance as a function of letter duration. Either no, two, or four additional

non-target objects appeared before the saccade. In Experiment 1, presaccadic non-targets were

task-irrelevant and had no effects on postsaccadic processing speed. In Experiment 2,

presaccadic non-targets were task-relevant because participants matched them against a probe

at trial end. Here, postsaccadic processing speed decreased with increasing number of

presaccadic non-targets. These findings show that objects compete for recognition across

saccades, but only if they are task-relevant. This reveals an attentional mechanism of task-

driven object recognition that is interlaced with active saccade-mediated vision (Schneider,

2013; Poth, Herwig, & Schneider, 2015).

46   Attentional Competition across Saccadic Eye Movements

47   Human goal-directed behavior heavily relies on the ability to recognize objects in the

48   environment visually. The capacity for visual object recognition, however, is severely limited

49   (for reviews, see Bundesen, Vangkilde, & Petersen, 2015; Duncan, 2006; Schneider, 1995).

50   Objects in the visual field must compete for limited attentional neuronal processing resources

51   (Desimone & Duncan, 1995). The more resources are allocated to an object, the faster it is

52   processed (Bundesen, 1990). Visual attention biases the resource allocation so that currently

53   important objects receive more resources than unimportant ones (Bundesen, 1990; Bundesen,

54   Habekost, & Kyllingsbaek, 2005). This directly impacts on object recognition. Only the first

55   few objects of a multi-object scene whose processing had been finished are encoded into a

56   limited-capacity visual working memory (VWM, also called visual short-term memory;

57   Bundesen, 1990; Bundesen et al., 2005, 2011). Once this has happened, the objects have been

58   recognized and become available for report and action (Bundesen, 1990; Schneider, 2013).

59   Importantly, the more equally relevant objects are present in the visual field, the smaller is

60   their share of neuronal resources and the slower are they processed (Bundesen, 1990;

61   Bundesen et al., 2005). Thus, visual object recognition increasingly suffers from attentional

62   competition between objects as more and more objects enter the visual field (Desimone &

63   Duncan, 1995).

64   There is extensive evidence that objects compete for object recognition within eye

65   fixations (e.g., Duncan, 2006; Poth, Petersen, Bundesen, & Schneider, 2014; Shibuya &

66   Bundesen, 1988; Vangkilde, Bundesen, & Coull, 2011), which are the periods in which the

67   eyes remain relatively stable (Findlay & Gilchrist, 2003; Land & Tatler, 2010). However, a

68   fundamental hallmark of human vision is thereby neglected: the active sampling of the visual

69   environment using rapid saccadic eye movements (for recent reviews, see Gegenfurtner,

70   2016; Rolfs, 2015; Schütz, Braun, & Gegenfurtner, 2011). Visual acuity is highest only at the

71   center of gaze, which falls on the central fovea of the eye's retina (e.g., Cowey & Rolls, 1974;

72   Curcio & Allen, 1990). Therefore, humans make saccadic eye movements that move the

73   fovea from one object to the next, so that the object is sampled in detail in the next fixation

74   (e.g., Findlay & Gilchrist, 2003; Land & Tatler, 2010). It is unclear whether attentional

75   competition between objects is constrained to a given eye fixation, or whether objects from

76   one fixation can compete with and thus impair the processing of objects in the next fixation.

77   One may hypothesize that there is no such transsaccadic attentional competition and

78   assume that successive eye fixations are entirely distinct visual processing episodes. This

79   *visual separation hypothesis* has intuitive appeal, because the retinal image is blurred and

80   visual information uptake is suppressed during saccades, which indeed separates one fixation

81   from the next (Krock & Moore, 2014; Wurtz, 2008). Moreover, only a limited number of

82   objects shown before a saccade can be reported after the saccade in accordance with a spatial

83   cue (Irwin, 1992; Irwin & Gordon, 1998). This has led to the proposal that only those objects

84   survive the saccade that are represented in limited-capacity VWM (or a similar transsaccadic

85   memory, respectively, for reviews, see Irwin, 1996; Mathôt & Theeuwes, 2011). The

86   competition between objects takes place before their encoding into VWM, and must hence

87   rely on object representations created prior to VWM encoding (Bundesen, 1990; Bundesen et

88   al., 2005). Therefore, if only representations in VWM survive the saccade, the competing

89   representations outside VWM should be lost across the saccade, so that there is no

90   transsaccadic competition.

91   What argues against the visual separation hypothesis is evidence that visual object

92   information outside VWM partially persists across the saccade (Irwin, 1992; Irwin, Brown, &

93   Sun, 1988). This persistence may be largely bound to the retinal locations of objects (Irwin et

94   al., 1988) which are moved by the saccade. However, because the competition for object

95   recognition concerns all visually available objects in the visual field (Bundesen, 1990; or at

96   least great parts thereof, Bundesen et al., 2005; Desimone & Duncan, 1995), this persisting

97   object information may compete with the actual objects in the next fixation. As a result, object

98 recognition in the next fixation should suffer per se from the object available in the previous

99 fixation.

100   Furthermore, it has been suggested that the competition for object recognition can

101 extend across changes and interruptions of visual input, such as those imposed by saccades,

102 but only if the objects are relevant to the task at hand (Schneider, 2013). This *task-driven*

103 *competition hypothesis* is directly based on Schneider's (2013) theory of "Task-dRiven visual

104 Attention and working Memory (TRAM)". According to TRAM, objects from the previous

105 fixation that are task-relevant but have not been fully processed will be encapsulated,

106 protected against being wiped-out by the saccade, and enter the competition for object

107 recognition in the next fixation. As a result, object recognition in this fixation should suffer

108 from all task-relevant objects of the previous fixation (except for those for which

109 correspondence between the fixations can be established, see also Poth, Herwig, & Schneider,

110 2015; Poth & Schneider, 2016b).

111   Here, we investigated the question of whether objects compete for object recognition

112 across saccadic eye movements. In two experiments, participants made saccades to peripheral

113 saccade target objects and then reported a letter that became visible within these objects after

114 the saccade. The letter was presented for a number of different durations and terminated by a

115 mask. We estimated the speed with which the letter was processed as the rate at which letter

116 report performance increased with increasing presentation duration (Bundesen, 1990, after a

117 minimum presentation duration has been exceeded). Importantly, visual processing speed of

118 the postsaccadic letter should directly reflect the amount of neuronal processing resources it

119 receives (Bundesen et al., 2005).

120   Experiment 1 investigated whether recognition of a postsaccadic object suffers from

121 attentional competition with presaccadic objects *per se*. To this end, the peripheral saccade

122 target appeared either alone, or was flanked by two, or four irrelevant non-target objects

123 (digits). The non-targets were extinguished as soon as participants made the saccade. Now, if

124   there was no attentional competition across the saccade, as per the proposed visual separation

125   hypothesis, then the number of presaccadic non-targets should have no effect on the visual

126   processing speed of the postsaccadic letter. In contrast, if there was attentional competition,

127   for instance due to lingering presaccadic representations (Irwin et al., 1988; Irwin, 1992), then

128   visual processing speed should be lower the more presaccadic non-targets are presented. To

129   preview the results, the number of presaccadic non-targets had no effect on the visual

130   processing speed of the postsaccadic letter, supporting the visual separation hypothesis.

131   Experiment 2 went on to test the task-driven competition hypothesis. It investigated

132   whether processing speed of a postsaccadic object suffers from attentional competition with

133   presaccadic objects when these are task-relevant. The paradigm was identical to the one of

134   Experiment 1, except that the non-targets were now task-relevant because they had to be

135   matched against a probe at trial end. The results of Experiment 2 support the task-driven

136   competition hypothesis: the more presaccadic non-targets had been shown, the lower was the

137   speed of processing the postsaccadic letter. Cross-experiment analyses corroborated that this

138   effect was indeed larger than the null effect of Experiment 1.

Method

Participants

141   Nine participants performed Experiment 1. An additional participant was excluded

142   from analysis due to letter reports at chance level (which prevented fitting the data with the

143   model below). Participants were between 22 and 30 years old ($MD = 25$ years), three were

144   male, six female, eight were right-, one was left-handed.

145   Eight different participants performed Experiment 2. An additional participant aborted

146   the experiment. Participants were between 20 and 31 years old ($MD = 23.5$ years), three were

147   male, five female, seven were right-, one was left-handed.

148   All participants reported normal or corrected-to-normal visual acuity and normal color

149   vision. They gave written informed consent before participation. The experiments followed

150 the ethical guidelines of the German Psychological Association (DGPs) and were approved by

151 Bielefeld University's ethics committee.

Apparatus and Stimuli

153 Participants performed the experiments in a semi-lit room. A head- and a chin-rest

154 ensured that they viewed the computer screen (G90fB, ViewSonic, Brea, CA, USA) from a

155 distance of 71 cm. The screen had a resolution of 1024×768 pixels at physical dimensions of

156 36×27 cm, a refresh rate of 100 Hz, and was controlled by a GeForce GTX 970 graphics

157 card (driver version 344.48, Nvidia, Santa Clara, CA, USA). A video-based tower-mounted

158 eye tracker (Eyelink 1000, SR Research, Ottawa, Ontario, CA) recorded the behavior of

159 participants' right eyes at a sampling rate of 1000 Hz. Responses were collected using a

160 standard computer keyboard (with QWERTZ layout).

161 The luminances and colors of stimuli were measured using an i1 Pro (X-Rite, Grand

162 Rapids, MI, USA) and are reported as CIE Lxy-coordinates. Stimuli were shown against a

163 gray background (screen center: $L = 11.605$ cd/m^2, $x = 0.294$, $y = 0.307$, averaged across left

164 and right stimulus locations: $L = 10.829$, $x = 0.288$, $y = 0.303$). A "+"-character ($0.34° \times$

165 $0.34°$ [degrees of visual angle], $L = 55.255$ cd/m^2, $x = 0.291$, $y = 0.304$) was used as central

166 fixation cross. The saccade target was a red circle appearing left or right of screen center (see

167 the procedure, $L = 29.886$ cd/m^2, $x = 0.606$, $y = 0.332$; measurements were averaged across

168 the two locations, as for all stimuli that appeared left and right of screen center). Non-targets

169 were the digits from 1 to 9 ($0.45° \times 0.93°$, $L = 21.855$ cd/m^2, $x = 0.535$, $y = 0.327$). Letter

170 stimuli were uppercase letters ($0.39° \times 0.67°$) from the set [ABDEFGHJKLMNOPRSTVXZ],

171 written in Arial font and in the background gray. For each individual participant, 30 pattern

172 masks were algorithmically created (as described in Poth et al., 2015; Poth & Schneider,

173 2016b) by overlaying nine black ($L = 0.139$ cd/m^2, $x = 0.252$, $y = 0.355$) mirror-reversed and

174 upside down letters (drawn randomly without replacement from the letter set) with horizontal

175   and vertical offsets. The large number of masks was intended to prevent adaptation to them.
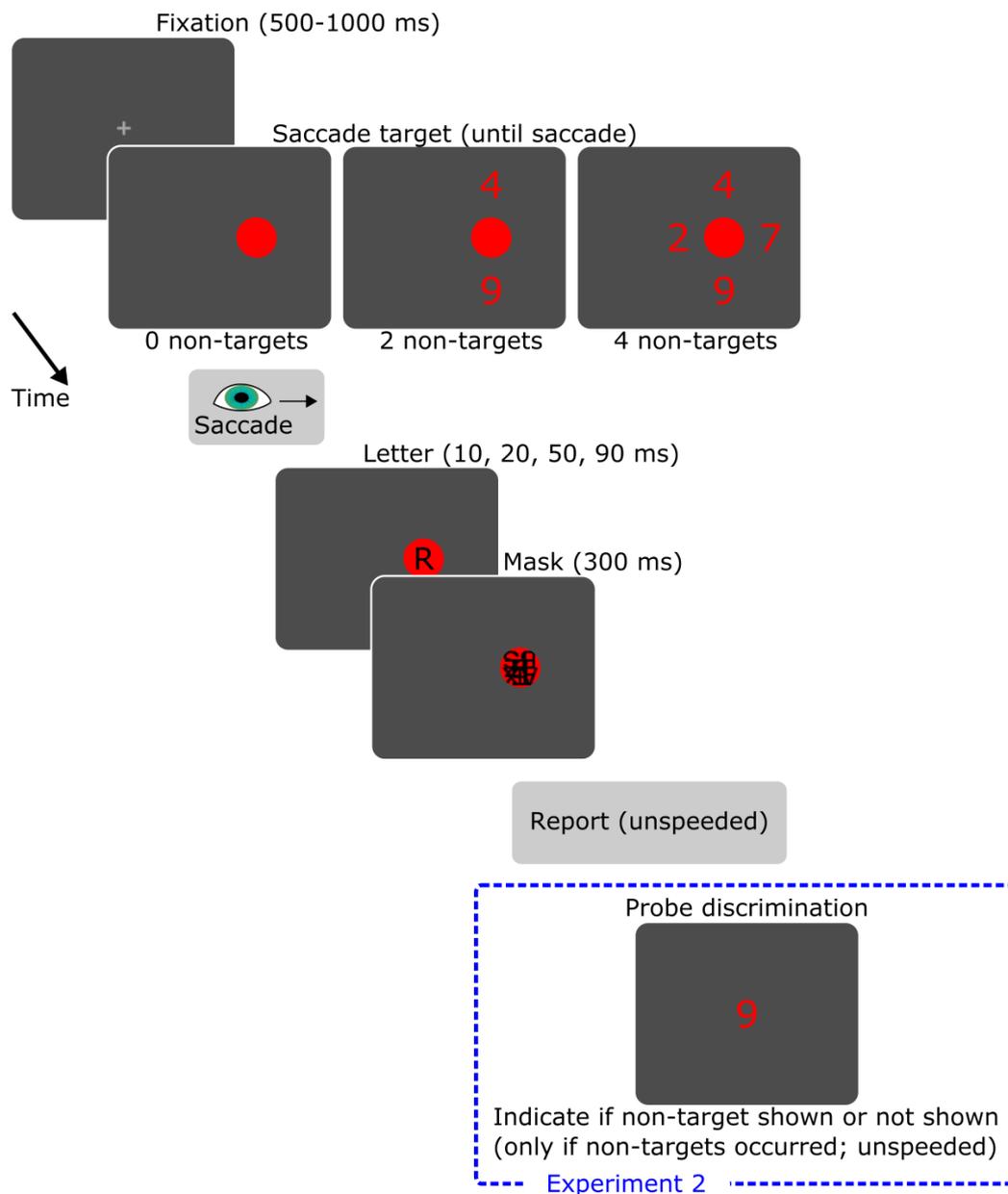
176   The pattern masks were placed on the red circle of the saccade target (see Figure 1).

177

178   **Procedure and Design**

179   Figure 1 illustrates the experimental paradigms of Experiment 1 and 2.

180

181

Figure 1. Paradigm of Experiment 1 and Experiment 2. At the beginning of a trial,

participants fixated a central fixation cross, which was followed by a red filled circle as

saccade target. The saccade target either appeared alone, was flanked by two, or by four digit

non-targets. When participants made a saccade to the peripheral target, any non-targets were

extinguished, and a letter appeared postsaccadically within the saccade target. The letter was

presented for different durations and terminated by a pattern mask. At the end of a trial,

participants reported the letter without time limit. Only in Experiment 2, a probe digit

followed participants' letter report and they indicated whether it was shown as a non-target in

the trial.

191   **Experiment 1.** Participants started each trial by pressing the space-bar. At the

192   beginning of a trial, participants fixated the central fixation cross for a random duration

193   between 500 and 1000 ms (that was a multiple of 10 ms). If participants moved gaze 2° from

194   the fixation cross in this interval, the trial was aborted and repeated at a randomly chosen

195   position in the sequence of remaining trials. The saccade target, the red circle, was then

196   shown at 7° to the left or right of screen center until participants made a saccade to it. In the *0*

197   *non-targets* condition, the saccade target was shown alone. In the *2 non-targets* condition,

198   task-irrelevant digits were shown above and one below the saccade target (both 2° from the

199   saccade target; see Figure 1). In the *4 non-targets* condition, digits were shown above, below,

200   left, and right of the saccade target (all in a distance of 2°; see Figure 1). Next, participants

201   made a saccade to the saccade target (detected as gaze position crossing a boundary of 3°

202   from screen center). If participants made no saccade until 400 ms after the onset of the

203   saccade target, the trial was aborted and repeated at a randomly chosen position in the

204   sequence of remaining trials. At the next screen refresh after saccade detection, any non-

205   targets were extinguished and a letter was shown within the saccade target. The letter was

206   randomly chosen from the letter set. The letter was presented for 10, 20, 50, or 90 ms. To

207   prevent an internal prolongation of letter presentation duration by early sensory (iconic)

208   memory (e.g., Irwin & Thomas, 2008), the letter was terminated by a pattern mask lasting for

209   300 ms (chosen randomly from the set of used masks). At the end of a trial, participants

210   reported the letter without time limit by using the keyboard. Participants had to type in a letter

211   to proceed, so that they had to guess if they were uncertain about the presented letter.

212   Participants performed 720 trials (plus repeated trials) in randomized order (30 trials ×

213   3 non-target conditions × 4 letter presentation durations × 2 locations of the saccade target).

214   Before the experiment, participants received instructions written on the screen, reported them

215   to the experimenter (who clarified any misunderstandings), and performed 40 training trials

216   (randomly drawn from the trials of the main experiment). A nine-point grid calibration of the

217   eye tracker was performed at the beginning of the experiment, after the training trials, and

218   after the 240th and 480th trial (irrespective of trial repetitions).

219   **Experiment 2.** The procedure and design were identical to Experiment 1 except that

220   any non-target digits flanking the saccade target were now task-relevant and had to be

221   remembered until the end of a trial. In the 2 and 4 non-targets condition, participants' report

222   of the postsaccadic letter was followed by a centrally displayed probe digit. Participants then

223   indicated whether the probe matched one of the non-targets shown on this trial (by pressing

224   "1" if they did not match and "0" if they did match; labels stating "Ja", German for "yes", and

225   "Nein", German for "no" were placed above these keys). The probe matched a non-target on

226   half of the trials per each cell of the design (see the design of Experiment 1).

227   **Eye movement analysis and trial exclusion criteria**

228   For all analyses, saccades were detected offline using the Eyelink 1000's algorithm

229   with a velocity threshold of $35° \times s^{-1}$ and an acceleration threshold of $9500° \times s^{-2}$. Trials were

230   excluded from analysis, if no saccade to the saccade target was made (i.e. participants crossed

231   the boundary for online saccade detection without meeting the offline saccade detection

232   criteria), if saccade latency (the time from the onset of the saccade target to saccade onset

233   detection) was below 100 ms, indicating anticipatory saccades, or if saccades missed the

234   saccade target location by more than 2°.

235   **Statistical analysis, estimation of visual processing speed and perceptual threshold**

236   Experimental conditions were compared using repeated measures analyses of variance

237   (ANOVAs, with type II sums-of-squares and η_G^2, as effect size, Bakeman, 2005), followed-up

238   by post-hoc paired t-tests (two-sided and with Cohen's d_z as effect size, Cohen, 1988). If the

239   t-test's assumption of normally distributed differences was violated, Wilcoxon-signed rank

240   tests were performed instead (for which we report r as effect size and a Z-value that we call

241   Z_w as test statistic and that was computed using the coin package for R, Hothorn, Hornik, van

242   de Wiel, & Zeileis, 2008).

243   To estimate visual processing speed and perceptual threshold, each participants' letter

244   report performance (the probability of reporting correctly) in each experimental condition was

245   assessed as a function of letter presentation duration and fit with the following model that is a

246   key ingredient of the "Theory of Visual Attention (TVA; Bundesen, 1990, see, also Bundesen

247   & Habekost, 2008):

248   $$p(t) = 1 - e^{-v\,(t\,-t_0)} + e^{-v(t\,-t_0)} \cdot p_{chance},$$

249   where p(t) is the probability that the letter has been recognized at time t and p_{chance} is the

250   probability of reporting correctly by chance (.05 in the current experiments). t_0 is the

251   perceptual threshold, the maximum ineffective presentation duration that has to be exceeded

252   for above-chance performance (see also Figure 2a). For times $t > t_0$, p(t) is exponentially

253   distributed with rate parameter v (the slope of the curve at t_0, see Figure 2a), which provides

254   the visual processing speed (recognized letters per unit time). Fitting was performed using a

255   maximum-likelihood procedure (by means of the optim function implemented in R (R Core

256   Team, 2016).

257   **Results**

258   Table 1 provides descriptive statistics of the dependent variables of both experiments.

259

260
261
262
263
264
265
266
267
268
269 Table 1 Descriptive Statistics of Experiment 1 and 2

	Processing speed		Perceptual threshold		Median saccade latency		Mean landing error		Probe discrimination performance	
N non-targets	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Experiment 1										
0	18.54	10.16	11.42	1.58	135	12	0.70	0.15	-	-
2	17.79	11.77	12.28	2.57	140	12	0.74	0.17	-	-
4	17.53	10.43	11.66	1.46	140	11	0.81	0.19	-	-
Experiment 2										
0	30.38	10.61	10.21	0.89	128	11	0.85	0.19	-	-
2	19.13	11.21	12.04	1.79	134	14	0.83	0.18	.63	.10
4	14.79	8.23	11.22	1.32	134	14	0.89	0.20	.57	.08

270 *Note*. Means (*M*) and standard deviations (*SD*) of participants' visual processing speed

271 (letters/s), perceptual threshold (ms), their probe discrimination performance (proportion

272 correct, only Experiment 2), their median saccade latency (ms), and mean saccade landing

273 error (°) in the three non-target conditions.

274

275   **Experiment 1**

276   **Excluded trials.** 0.05% of the trials were excluded because no saccade was made,

277   3.41% because saccade latency fell short of 100 ms, and 1.11% because saccades missed their

278   target by more than 2°.

279   **Visual processing speed and perceptual threshold.** The model fit the data well, so

280   that the correlation between predicted and observed values on average exceeded .91 ($SD =$

281   .05) in all three non-target conditions. Visual processing speed did not differ significantly

282   between the conditions, see Figure 2b, $F(2, 16) = 0.523$, $p = .603$, $\eta_G^2 = .002$. To assess

283   whether it was more likely that there was no difference between the three conditions, the

284   Bayes Factor in favor of the null hypothesis (BF_{01}, with participant as random factor and

285   whereby values larger 1 supporting the null hypothesis) was computed using the Bayes Factor

286   package for R (Morey & Rouder, 2015). According to this analysis, it was more likely that the

287   visual processing speed did not differ between the three conditions, $BF_{01} = 3.275$. Likewise,

288   there were no significant differences between the perceptual thresholds in the three

289   conditions, see Figure 2c, $F(2, 16) = 0.733$, $p = .496$, $\eta_G^2 = .038$ and also here it was more

290   likely that such differences were absent, $BF_{01} = 2.746$.
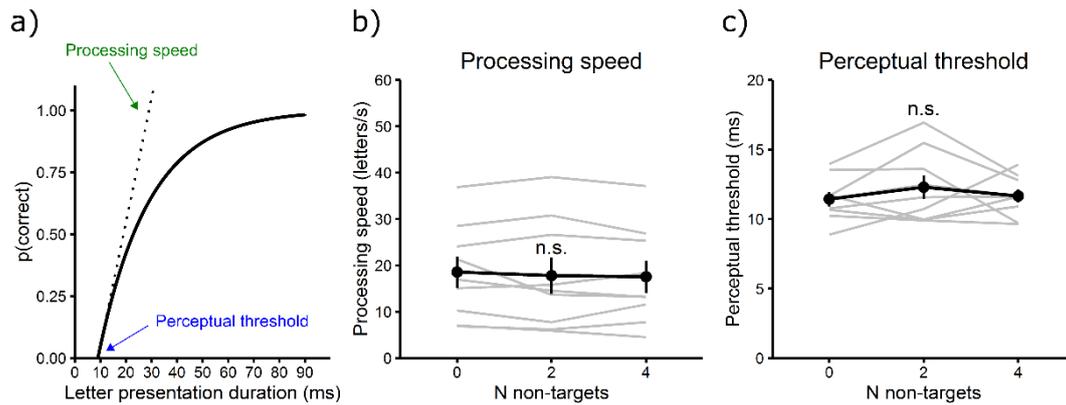
291

292

293

294

295

296

297

Figure 2.[a]In both experiments, letter report performance was assessed as a function of letter

presentation duration and fit with an exponential model (Bundesen, 1990), yielding the

perceptual threshold (x-intercept, maximum ineffective presentation duration) and processing

speed (exponential rate and, equivalently, slope of the curve at the perceptual threshold). Note

that the model actually used for data analysis allowed for chance performance in addition.

[b,c]Gray lines represent individual participants, black lines and circles represent sample means.

[b]Visual processing speed (letters/s) and [c]the perceptual threshold (ms) did not differ

significantly between the different numbers of non-targets. Error-bars represent ± 1 standard

error.

307   **Saccade latencies and landing errors.** In each non-target condition, participants'

308   saccade latency was assessed as the median saccade latency across trials. Saccade latency

309   differed significantly between the non-target conditions, $F(2, 16) = 21.921$, $p < .001$, $\eta_G^2 =$

310   .046 (cf. Table 1). Specifically, saccades were initiated significantly faster in the 0 non-targets

311   condition than in the 2 non-targets condition, $t(8) = -5.498$, $p < .001$, $d_z = -1.83$ and the 4 non-

312   targets condition, $t(8) = -5.393$, $p < .001$, $d_z = -1.80$. However, saccade latency did not differ

313   significantly between the 2 and the 4 non-targets conditions, $t(8) = -0.555$, $p = .594$, $d_z = -$

314   0.185.

315   Saccade landing error in the three non-target conditions was assessed as participants'

316   mean Euclidian distance (in °) between saccade landing sites and saccade target locations

317   across trials. Saccade landing error differed significantly between non-target conditions, $F(2,$

318     $16) = 6.941$, $p = .007$, $\eta_G^2 = .073$ (cf. Table 1). Saccades were significantly more accurate in

319   the 0 non-targets condition compared with the 4 non-targets condition, $t(8) = -3.486$, $p = .008$,

320     $d_z = -1.16$. The difference in saccade landing error between the 0 and the 2 non-targets

321   condition was not significant (but might be seen as close to for this test that was uncorrected

322   for multiple comparisons), $t(8) = -2.025$, $p = .078$, $d_z = -0.67$, and the same was the case for

323   the difference between the 2 and 4 non-targets conditions, $t(8) = -1.973$, $p = .084$, $d_z = -0.66$.

324   Moreover, the non-target conditions did not differ significantly with respect to participants'

325   saccade precision (their *SD* of saccade landing errors across trials), $F(2, 16) = 0.930$, $p = .415$,

326     $\eta_G^2 = .009$.

327   **Experiment 2**

328   **Excluded trials.** 0.07 % of the trials were excluded from analysis because no saccade

329   was made, 8.68 % because saccade latency was below 100 ms, and 6.22% because saccades

330   missed the target by more than 2°.

331   **Visual processing speed and perceptual threshold.** The exponential model fit

332   participants' data well so that the mean correlation of predicted and observed values exceeded

333   .94 (*SD* = .04) in each non-target condition. In contrast to Experiment 1, visual processing

334   speed differed significantly between the three non-target conditions, see Figure 3a, $F(2, 14) =$

335   8.750, p = .003, η_G^2 = .326, BF_{01} = 0.050 Visual processing speed was significantly higher in

336   the 0 non-targets condition than in the 2 non-targets condition, Z_w = 2.381, p = .016, r = .842,

337   and in the 4 non-targets condition, Z_w = 2.521, p = .008, r = .891. Likewise, processing speed

338   was significantly higher in the 2 compared with the 4 non-targets condition, Z_w = 2.240, p =

339   .023, r = .792 (though the difference was only close to significance after Bonferroni
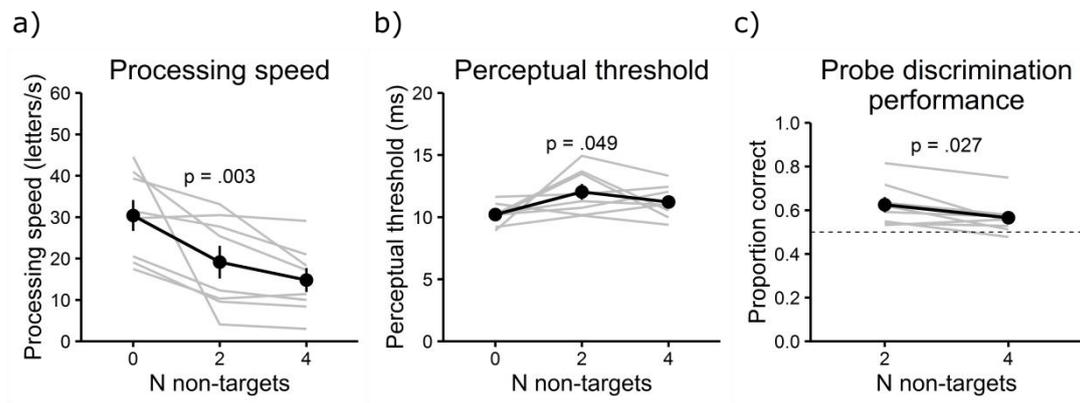
340   correction, p = .070).

341

342

343

344

345

346

347

Figure 3. Gray lines represent individual participants, black lines and circles represent sample

means. [a]Visual processing speed (letters/s) and [b]the perceptual threshold differed significantly

between the non-target conditions. [c]Participants performed above chance level (dashed line)

in probe discrimination, whereby performance was higher in the 2 than in the 4 non-targets

condition. Error-bars represent ± 1 standard error.

353

354   Moreover, a mixed ANOVA (with type III sums-of-squares) with experiment as

355   between-subjects factor showed that the differences in the effect of non-targets on visual

356   processing speed between the experiments were significant (i.e. there was a significant

357   interaction between experiment and condition), $F(2, 30) = 7.957$, p (Greenhouse-Geisser

358   corrected) $= .007$, $\varepsilon = 0.651$, $\eta_G^2 = .088$.

359   In Experiment 2, the perceptual threshold differed significantly between the three non-

360   target conditions as well, see Figure 3b, $F(2, 14) = 3.762$, $p = .049$, $\eta_G^2 = 0.250$, $BF_{01} = 0.343$.

361   It was close to significantly lower in the 0 non-targets condition than in the 2 non-targets

362   condition, $Z_w = -1.960$, $p = .055$, $r = -.693$, but significance disappeared after Bonferroni

363   correction ($p = .164$). Likewise, the perceptual threshold was also lower in the 0 compared

364   with the 4 non-targets condition, $Z_w = -2.100$, $p = .039$, $r = -0.74$, before but not after

365   Bonferroni correction ($p = .117$). The difference between the 2 and 4 non-targets conditions

366   was not significant, $Z_w = 0.980$, $p = .383$, $r = .347$. It is important to note that the differences

367   between the conditions in the perceptual threshold do not conflict with interpretation of the

368   differences in visual processing speed. This is evident from the fact that the differences in

369   perceptual threshold between the 0 and 2, and 0 and 4 non-targets conditions were in the

370   opposite direction of their difference in visual processing speed (see Table 1 and Figure 3).

371   **Probe discrimination performance.** One-sided one-sample t-tests showed that

372   participants' probe discrimination performance (proportion of correct responses) was above

373   chance in both, the 2 and 4 non-targets conditions, see Figure 3c, both ts(7) > 2.273, both ps <

374   .029. Moreover, probe discrimination was significantly higher in the 2 than the 4 non-targets

375   condition, $t(7) = 2.790$, $p = .027$, $d_z = 0.99$.

376   **Saccade latencies and landing errors.** As in Experiment 1, saccade latency differed

377   significantly between the non-target conditions, $F(2, 14) = 13.514$, $p < .001$, $\eta_G^2 = .053$.

378   Saccades were initiated faster in the 0 non-targets condition, than in the 2 non-targets, $t(7) = -

379   3.689$, $p = .008$, $d_z = -1.30$, and the 4 non-targets condition, $t(7) = -4.168$, $p = .004$, $d_z = -1.47$.

380   The 2 and 4 non-targets conditions did not differ significantly, $t(7) = -0.303$, $p = .770$, $d_z = -$

381   0.11. Note that the differences in saccade latency between the non-targets conditions should

382   not conflict with the differences in visual processing speed. The differences in saccade latency

383   were in the opposite direction of the differences in visual processing speed. This means that

384   higher processing speed did not come at a cost of slower saccades. Conversely, note that the

385   onset of the postsaccadic letter was contingent upon the saccade (i.e. on the time eye position

386   crossed a boundary). Therefore, saccades that were initiated faster should not have exerted

387   strong effects on the presentation duration of the postsaccadic letter, which against argues

388   against a conflict of differences in saccade latency and visual processing speed.

389   Saccade landing errors did not differ significantly between the three conditions, $F(2,$

390    $14) = 0.904$, $p = .427$, $\eta_G^2 = .017$, and the same was the case for saccade precision, $F(2, 14) =$

391   0.364, $p = .701$, $\eta_G^2 = .011$.

General Discussion

393   We asked whether attentional competition between objects for recognition is confined

394   to a single eye fixation or whether it extends across saccadic eye movements. More

395   specifically, we raised the question if presaccadic and postsaccadic objects compete such that

396   postsaccadic object recognition is slowed down as more objects are viewed before the

397   saccade. The results of Experiment 1 argue against this hypothesis. The number of task-

398   irrelevant presaccadic objects had no effect on the speed with which a postsaccadic object was

399   processed. As such, this finding may support the visual separation hypothesis, holding that

400   successive fixations are distinct processing episodes without any competitive interactions

401   between them. This hypothesis was, however, falsified by Experiment 2 which found

402   attentional competition across the saccade when presaccadic objects were relevant to the

403   current task. The speed of processing a postsaccadic object for its recognition here indeed

404   decreased with increasing number of presaccadic objects. Cross-experiment analyses

405   confirmed that this effect of transsaccadic attentional competition was indeed greater in when

406 presaccadic objects were task-relevant. In this way, the present results provide evidence for

407 the task-driven competition hypothesis, which maintains that attentional competition between

408 objects across saccades critically depends on the objects' task-relevance (Schneider, 2013).

409   **Task-driven attentional competition across saccadic eye movements**

410   The present findings implicate the current task as a driving force of object recognition

411 across saccadic eye movements. This is well in line with Schneider's (2013) TRAM theory,

412 which proposes a set of mechanisms that enable visual object recognition across saccades, as

413 well as other changes and interruptions of visual input. TRAM is based on Bundesen's (1990)

414 TVA theory. TVA assumes that visual processing for object recognition relies on categorizing

415 objects as having certain features. When a categorization of an object (e.g. that it is the letter

416 "R") has entered VWM, then storage space for all other categorizations of the object (e.g. that

417 it is red) is reserved, and eventually the object will be represented in VWM with all its

418 features. Critically, all object categorizations participate in a parallel and capacity-limited race

419 towards VWM. Those objects whose processing finishes first are encoded into VWM until the

420 number of objects that can maximally be retained in VWM is reached. The speed with which

421 an object categorization is processed is proportional to the share of all available neuronal

422 processing resources that is allocated to the object (Bundesen et al., 2005). Processing

423 resources are allocated to objects according to the objects' *attentional weights*, which indicate

424 the objects' current importance. More specifically, attentional weights code for the products

425 of the sensory evidence that the object has a feature and the task-relevance of this feature,

426 summed across all possible features. Because processing resources are limited, they are

427 allocated to objects in a normalized fashion, according to a given object's attentional weight

428 relative to the sum of attentional weights of all objects in the visual field. Thus, the higher the

429 relative attentional weight of an object, the more resources it receives and the faster its

430 categorizations race toward VWM. As a result, the higher are the object's chances of entering

431 VWM. Thus, attentional competition between objects for recognition finds expression in a

432   lower speed of processing object categorizations as more and more objects enter the visual

433   field.

434   TRAM extends TVA to the situation of visual processing across saccadic eye

435   movements. TRAM assumes that after encoding into VWM, objects are consolidated into a

436   *passive* VWM, which retains the objects for cognitive processing and action control for more

437   extended periods (e.g., across action steps, and without the capacity constraints of classical

438   VWM). The consolidation into passive VWM consumes the neuronal processing resources

439   that also accomplish encoding of the objects into VWM. If a saccade is made while

440   consolidation has not yet finished, then the attentional weights of these objects are

441   encapsulated, so that the associated neuronal resources cannot be redistributed to process new

442   objects. Therefore, fewer resources are available in the next fixation, which slows down

443   visual processing. One assumption of TRAM is most important for the current findings: For

444   the attentional weight of an object to be encapsulated, the object must be task-relevant. This

445   explains why postsaccadic object recognition was unaffected by irrelevant presaccadic objects

446   in Experiment 1 but slowed down by the task-relevant ones in Experiment 2. The irrelevant

447   presaccadic non-targets might have been encoded into VWM in Experiment 1 (cf. Bundesen,

448   1990), as there should have been enough space for at least some of them (this is supported by

449   the above-chance probe discrimination in Experiment 2, which should require prior encoding

450   of the to-be-discriminated non-targets into VWM, Poth & Schneider, 2016a). However,

451   because the objects were irrelevant to the task, their attentional weights and associated

452   processing resources should not have been encapsulated when the saccade was made. Instead,

453   their attentional weights should have been cleared, freeing neuronal resources for unimpaired

454   processing in the next fixation. In contrast, in Experiment 2 the attentional weights of the

455   presaccadic objects should have been encapsulated because the objects were task-relevant and

456   because the time before the saccade did not suffice to complete consolidation into passive

457   VWM (see, e.g., Carbone & Schneider, 2010). This should have cut the resources available

458   for processing the postsaccadic object and lead to the observed decrement in visual processing

459   speed for this object.

Object correspondence could prevent ubiquitous attentional competition

461   The current experiments investigated the effects of presaccadic objects that were

462   extinguished across the saccade on the speed with which a postsaccadic object was processed.

463   However, one may ask how postsaccadic processing is affected if objects remain present

464   across the saccade. Would there be attentional competition between the presaccadic and the

465   postsaccadic instance of the same object? This would mean that every saccade imposed costs

466   on visual processing, which would impair object recognition in most situations of human life.

467   TRAM (Schneider, 2013) proposes a solution to this problem. After every saccade, the

468   visual system tests whether a currently viewed object corresponds to one viewed before the

469   saccade. If this test for object correspondence is positive, the postsaccadic object is processed

470   using the same attentional weight as the presaccadic one and the features of the postsaccadic

471   object update the representation of the presaccadic object which is now in VWM. In contrast,

472   if the test for transsaccadic object correspondence is negative, then the presaccadic attentional

473   weight of the object in VWM is encapsulated and a new representation is created for the

474   postsaccadic object. As a consequence, processing of the postsaccadic object is impaired

475   because the encapsulated attentional weight leaves fewer processing resources available.

476   Consistent with this hypothesis, breaking object correspondence by blanking the postsaccadic

477   object briefly, or by changing an object's contrast-polarity (Poth et al., 2015) or color (Poth

478   & Schneider, 2016b) across the saccade deteriorates postsaccadic object recognition. Thus,

479   taken together, an ubiquitous attentional competition between objects across every saccade

480   may be prevented by an additional mechanism that establishes transsaccadic object

481   correspondence (Schneider, 2013).

482

483

484   **Conclusion**

485   It has long been clear that within an eye fixation, objects compete for limited neuronal

486   processing resources that are necessary for object recognition (Bundesen, 1990; Desimone

487   & Duncan, 1995; Duncan & Humphreys, 1989). Our findings reveal that such a competition

488   for object recognition also takes place across the saccade, but only if the objects are relevant

489   to the current task. Within eye fixations, processing resources are allocated to objects

490   according to their current task-relevance, supporting recognition of the most important ones

491   (e.g., Bundesen, 1990). Across saccades, resources remain allocated to objects but, again,

492   only if the objects are currently important. If they are not important, resources are freed so

493   that new objects can be processed faster in the next fixation (Schneider, 2013). Taken

494   together, we conclude that there must be mechanisms for task-driven object recognition that

495   are not restricted to eye fixations but intertwined with active saccade-mediated vision.

496

497

498

References

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*(3), 379–384. doi:10.3758/BF03192707

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523–547. doi:10.1037/0033-295X.97.4.523

Bundesen, C., Habekost, T., & Kyllingsbaek, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review*, *112*(2), 291–328. doi:10.1037/0033-295X.112.2.291

Bundesen, C., & Habekost, T. (2008). *Principles of Visual Attention: Linking Mind and Brain.* Oxford, UK: Oxford University Press.

Bundesen, C., Habekost, T., & Kyllingsbaek, S. (2011). A neural theory of visual attention and short-term memory (NTVA). *Neuropsychologia*, *49*(6), 1446–1457. doi:10.1016/j.neuropsychologia.2010.12.006

Bundesen, C., Vangkilde, S., & Petersen, A. (2015). Recent developments in a computational theory of visual attention (TVA). *Vision Research*, *116*, 210–218. doi:10.1016/j.visres.2014.11.005

Carbone, E., & Schneider, W. X. (2010). The control of stimulus-driven saccades is subject not to central, but to visual attention limitations. *Attention, Perception & Psychophysics*, *72*(8), 2168–2175. doi:10.3758/APP.72.8.2168

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). New York, NY, USA: Psychology Press.

Cowey, A., & Rolls, E. T. (1974). Human cortical magnification factor and its relation to visual acuity. *Experimental Brain Research*, *21*(5). doi:10.1007/BF00237163

Curcio, C. A., & Allen, K. A. (1990). Topography of ganglion cells in human retina. *The Journal of Comparative Neurology*, *300*(1), 5–25. doi:10.1002/cne.903000103

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222. doi:10.1146/annurev.ne.18.030195.001205

Duncan, J. (2006). EPS Mid-Career Award 2004: Brain mechanisms of attention. *Quarterly Journal of Experimental Psychology*, *59*(1), 2–27. doi:10.1080/17470210500260674

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458. doi:10.1037/0033-295X.96.3.433

Findlay, J. M., & Gilchrist, I. D. (2003). *Active Vision: The Psychology of Looking and Seeing*. New York: Oxford University Press.

Gegenfurtner, K. R. (2016). The interaction between vision and eye movements. *Perception.* doi:10.1177/0301006616657097

Hothorn, T., Hornik, K., van de Wiel, M. A., & Zeileis, A. (2008). Implementing a Class of Permutation Tests: The coin Package. *Journal of Statistical Software*, *28*(8), 1–23.

Irwin, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 307–317. doi:10.1037/0278-7393.18.2.307

Irwin, D. E. (1996). Integrating information across saccadic eye movements. *Current Directions in Psychological Science*, *5*(3), 94–100. doi:10.1111/1467-8721.ep10772833

Irwin, D. E., Brown, J. S., & Sun, J.-s. (1988). Visual masking and visual integration across saccadic eye movements. *Journal of Experimental Psychology: General*, *117*(3), 276–287. doi:10.1037/0096-3445.117.3.276

Irwin, D. E., & Gordon, R. D. (1998). Eye movements, attention and trans-saccadic memory. *Visual Cognition*, *5*(1-2), 127–155. doi:10.1080/713756783

Irwin, D. E., & Thomas, L. E. (2008). Visual sensory memory. In S. J. Luck & A. Hollingworth (Eds.), *Visual Memory* (pp. 9–41). New York: Oxford University Press.

Krock, R. M., & Moore, T. (2014). The influence of gaze control on visual perception: Eye movements and visual stability. *Cold Spring Harbor Symposia on Quantitative Biology*, *79*, 123–130. doi:10.1101/sqb.2014.79.024836

Land, M. F., & Tatler, B. W. (2010). *Looking and Acting: Vision and Eye Movements in Natural Behaviour*. Oxford, UK: Oxford University Press.

Mathôt, S., & Theeuwes, J. (2011). Visual attention and stability. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1564), 516–527. doi:10.1098/rstb.2010.0187

Morey, R. D., & Rouder, J. N. (2015). BayesFactor: Computation of Bayes Factors for Common Designs. Retrieved from https://CRAN.R-project.org/package=BayesFactor

Poth, C. H., Herwig, A., & Schneider, W. X. (2015). Breaking object correspondence across saccadic eye movements deteriorates object recognition. *Frontiers in Systems Neuroscience*, *9*, 176. doi:10.3389/fnsys.2015.00176

Poth, C. H., Petersen, A., Bundesen, C., & Schneider, W. X. (2014). Effects of monitoring for visual events on distinct components of attention. *Frontiers in Psychology*, *5*, 930. doi:10.3389/fpsyg.2014.00930

Poth, C. H., & Schneider, W. X. (2016a). Episodic short-term recognition requires encoding into visual working memory: Evidence from probe recognition after letter report. *Frontiers in Psychology*, *7*, 1440. doi:10.3389/fpsyg.2016.01440

Poth, C. H., & Schneider, W. X. (2016b). Breaking object correspondence across saccades impairs object recognition: The role of color and luminance. *Journal of Vision*, *16*(11), 1. doi:10.1167/16.11.1

R Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rolfs, M. (2015). Attention in Active Vision: A Perspective on Perceptual Continuity Across Saccades. *Perception*, *44*(8-9), 900–919. doi:10.1177/0301006615594965

Schneider, W. X. (1995). VAM: A neuro-cognitive model for visual attention control of segmentation, object recognition, and space-based motor action. *Visual Cognition*, *2*(2-3), 331–376. doi:10.1080/13506289508401737

Schneider, W. X. (2013). Selective visual processing across competition episodes: a theory of task-driven visual attention and working memory. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(1628), 20130060. doi:10.1098/rstb.2013.0060

Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Eye movements and perception: A selective review. *Journal of Vision*, *11*(5). doi:10.1167/11.5.9

Shibuya, H., & Bundesen, C. (1988). Visual selection from multielement displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(4), 591–600. doi:10.1037/0096-1523.14.4.591

Vangkilde, S., Bundesen, C., & Coull, J. T. (2011). Prompt but inefficient: Nicotine differentially modulates discrete components of attention. *Psychopharmacology*, *218*(4), 667–680. doi:10.1007/s00213-011-2361-x

Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Research*, *48*(20), 2070–2089. doi:10.1016/j.visres.2008.03.021

Study 4

Episodic short-term recognition requires encoding into visual working memory:
Evidence from probe recognition after letter report

Episodic Short-Term Recognition Requires Encoding into Visual Working Memory: Evidence from Probe Recognition after Letter Report

Christian H. Poth* and Werner X. Schneider

Neuro-Cognitive Psychology, Department of Psychology and Cluster of Excellence Cognitive Interaction Technology, Bielefeld University, Bielefeld, Germany

Human vision is organized in discrete processing episodes (e.g., eye fixations or task-steps). Object information must be transmitted across episodes to enable episodic short-term recognition: recognizing whether a current object has been seen in a previous episode. We ask whether episodic short-term recognition presupposes that objects have been encoded into capacity-limited visual working memory (VWM), which retains visual information for report. Alternatively, it could rely on the activation of visual features or categories that occurs before encoding into VWM. We assessed the dependence of episodic short-term recognition on VWM by a new paradigm combining letter report and probe recognition. Participants viewed displays of 10 letters and reported as many as possible after a retention interval (whole report). Next, participants viewed a probe letter and indicated whether it had been one of the 10 letters (probe recognition). In Experiment 1, probe recognition was more accurate for letters that had been encoded into VWM (reported letters) compared with non-encoded letters (non-reported letters). Interestingly, those letters that participants reported in their whole report had been near to one another within the letter displays. This suggests that the encoding into VWM proceeded in a spatially clustered manner. In Experiment 2, participants reported only one of 10 letters (partial report) and probes either referred to this letter, to letters that had been near to it, or far from it. Probe recognition was more accurate for near than for far letters, although none of these letters had to be reported. These findings indicate that episodic short-term recognition is constrained to a small number of simultaneously presented objects that have been encoded into VWM.

Keywords: visual working memory, visual attention, episodic memory, object recognition, short-term memory

INTRODUCTION

Visual information processing is organized in discrete episodes. This is most evident from the fact that the uptake of visual information is largely limited to eye fixations, discrete periods of stable eye position that are interrupted by fast saccadic eye movements (e.g., Krock and Moore, 2015). However, on a greater time scale, processing episodes can also be defined by steps of sensorimotor

actions, other task-demands, and changes in the visual environment (Petersen et al., 2012; Duncan, 2013; Schneider, 2013; Herwig, 2015; Poth et al., 2015; Poth and Schneider, 2016). To remain oriented in time and space and to act guided by vision, visual information from consecutive processing episodes must be linked. This is particularly evident from tasks requiring to recognize that objects (or subjects) have been viewed recently (e.g., Sternberg, 1966; Wickelgren, 1970; Kahana and Sekuler, 2002; Zhou et al., 2004; Donkin and Nosofsky, 2012). For example, imagine you are standing at a busy inner-city intersection and someone shows you a picture of a dog that just went missing and asks if you have seen it. To answer this question, you must be able to recognize if the dog appeared in one of the many recent processing episodes that consisted of your eye fixations, steps of your actions, and periods of cars passing by. Such tasks require *episodic short-term recognition*: the cognitive function of recognizing whether a now-present object has been contained in a recently passed visual processing episode[1] (cf. Kahana and Sekuler, 2002; Zhou et al., 2004; Donkin and Nosofsky, 2012).

How is episodic short-term recognition accomplished? What are its underlying mechanisms? First of all, to recognize that an object has been present before, the object must be represented internally. Several views on visual processing posit that initially, objects are represented by activating their corresponding feature or category representations in visual long-term memory (Cowan, 1988; Bundesen, 1990; Henderson, 1994; Henderson and Anes, 1994; Eriksson et al., 2015; cf. Oberauer, 2002; LaRocque et al., 2014; for a more general overview, see Palmeri and Tarr, 2008). These representations code for visual features and categories of objects that have been acquired through past visual experience and are often called *visual types* (e.g., Kanwisher, 1987; Kahneman et al., 1992; although other terms are in use as well, e.g., Duncan and Humphreys, 1989; Bundesen, 1990). Visual types represent objects in a multidimensional feature and category space and they may also represent exemplars of certain objects (cf. Kahana and Sekuler, 2002; Nosofsky et al., 2011; Donkin and Nosofsky, 2012).

Critically, activating an object's visual type (feature, category) is only considered an initial step of processing (Duncan and Humphreys, 1989; Bundesen, 1990; Bundesen et al., 2005; Kyllingsbæk, 2014). This activation does neither suffice to act upon the object nor to consciously perceive the object in the sense that it can be reported. Importantly, the activation is "pre-attentive" in the sense of being unselective: it proceeds likewise for all objects in the visual field (or parts of the visual field, depending on pre-existing spatial biases, Bundesen and Habekost, 2008, p. 117, and retinal inhomogeneity, Strasburger et al., 2011). That is, it proceeds before mechanisms of visual attention select task-relevant objects for further processing at the expense of task-irrelevant ones (e.g., Duncan and Humphreys, 1989; Bundesen,

1990; Bundesen et al., 2005; Duncan, 2006; Poth et al., 2014). For action and report, objects must be attended, processed further, and eventually encoded into visual working memory (VWM; Duncan and Humphreys, 1989; Bundesen, 1990; Cowan, 2001; Bundesen et al., 2005; Schneider, 2013; note that we use VWM synonymously to the also common term of visual short-term memory).

Visual working memory consists of a mechanism for retaining visual object representations accessible over short time-windows (for reviews, see Luck, 2008; Bundesen et al., 2011; Luck and Vogel, 2013; LaRocque et al., 2014; Ma et al., 2014). In this way, VWM may provide an essential basis for further processing these representations, as recoding them into other representational formats (e.g., the verbal format) so that they can be retained and used by non-visual mechanisms of working memory (e.g., Logie, 2011). The capacity of VWM is limited so that it can only hold about three to four objects (e.g., Sperling, 1960; Shibuya and Bundesen, 1988; Luck and Vogel, 1997; Dyrholm et al., 2011; Poth et al., 2014; note that capacity is also limited in the number of object features, Wheeler and Treisman, 2002; Oberauer and Eichenberger, 2013, and the precision of object features, Wilken and Ma, 2004; Bays and Husain, 2008). Which of all available objects are encoded into VWM depends on selection by visual attention (e.g., Duncan and Humphreys, 1989; Bundesen, 1990; Bundesen et al., 2005; Duncan, 2006; Poth et al., 2014). Because of the limited capacity of VWM, all visually available objects may initially and (pre-attentively) activate visual types in visual long-term memory, but only a limited number of objects is (attentively) processed up to the level of VWM (Duncan and Humphreys, 1989; Bundesen, 1990; Bundesen et al., 2005). Encoding objects into VWM is a core requirement of visually controlled behavior, because objects can only be reported and used for action when they are represented in VWM (Duncan and Humphreys, 1989; Bundesen, 1990; Bundesen et al., 2005). This paper focuses on the open question of whether encoding into VWM is also necessary for episodic short-term recognition.

Episodic short-term recognition requires comparisons of object representations of a recently preceding processing episode with representations of objects of the current episode. This can be conceptualized as a decision process (e.g., Pearson et al., 2014) which is driven by the degree of similarity between these two kinds of representations (e.g., Ratcliff, 1978; Donkin and Nosofsky, 2012; cf. Kahana and Sekuler, 2002). Two rival hypotheses can be advanced regarding the role of VWM in this comparison process (based on the literature covered above). According to the *VWM-encoding* hypothesis, episodic short-term recognition of an object from a previous episode requires that the object has been encoded into VWM. Consequently, objects that have not been processed up to the level of VWM cannot be used for episodic short-term recognition. Alternatively, the *type-activation* hypothesis states that episodic short-term recognition is also possible for objects which have not been encoded into VWM but whose mere presentation has activated their visual types in visual long-term memory. This means that episodic short-term recognition is possible for all external objects that have been visually available within recent eye fixations. In such a case, activations of visual types could extend into the

---

[1]Note that the term episodic short-term recognition refers to the described cognitive function (in the sense of a cognitive task-requirement). In this way, the concept of episodic short-term recognition does not include any assumptions about the cognitive mechanisms enabling to fulfill this function (such as for example, interacting mechanisms of episodic long-term memory or working memory).

next processing episode. These remaining activations could be matched against activations elicited by objects of this episode. A resulting signal could then allow the comparison of object representations from the previous episode and from the actual environment underlying episodic short-term recognition (e.g., Ratcliff, 1978; Donkin and Nosofsky, 2012). Such a mechanism could be similar to mechanisms assumed to produce attention-independent priming effects, where the presentation of objects facilitates their subsequent object recognition (e.g., Kahneman et al., 1992; Henderson, 1994; Henderson and Anes, 1994; Jensen and Lisman, 1998) or affects motor responses to other stimuli (even if the objects are not discriminable, Klotz and Neumann, 1999, and hence not in VWM, Bundesen, 1990).

Here, we aimed at deciding between the two hypotheses. In two experiments, we asked whether episodic short-term recognition of an object requires that this object has previously been encoded into capacity-limited VWM. To approach this question, we introduced a new paradigm combining letter report with probe recognition.

EXPERIMENT 1

In Experiment 1, participants performed a whole report task (e.g., Sperling, 1960; Shibuya and Bundesen, 1988) which was combined with a probe recognition task. They briefly viewed displays of to-be-memorized letters (memory letters) and then, after a retention interval, reported as many letters as they could. The retention interval outlasted early sensory memory (e.g., Sperling, 1960; Phillips, 1974; Irwin and Thomas, 2008) so that letter reports should have required retention in VWM (followed by a recoding into a verbal format on which the actual report was based, e.g., Logie, 2011; Baddeley, 2012). Memory letters were always 10 different ones, exceeding VWM capacity and thus ensuring participants could never report all letters (Sperling, 1960; Shibuya and Bundesen, 1988). After reporting the letters, a single probe letter appeared within the same trial and participants indicated whether or not the probe had been shown as one of the previous memory letters. Importantly, the probe was either one of the memory letters *and* reported (*reported condition*), or one of the memory letters but not reported (*non-reported condition*), or it was a letter not contained in the set of memory letters (*not shown condition*).

Here, episodic short-term recognition was assessed as performance in probe recognition, that is, in indicating whether or not the probe letter had been shown as one of the memory letters. Which memory letters were encoded into VWM was assessed by preceding letter reports. Since VWM is defined by the accessibility of its content (e.g., Bundesen, 1990; Bundesen et al., 2005; Schneider, 2013; but see, Soto et al., 2011), reported letters must have been in VWM by definition. Following a number of theories (e.g., Bundesen, 1990; Bundesen et al., 2005; Martens and Wyble, 2010; Schneider, 2013), we assume that letters which were not reported did not enter VWM. Consequently, the VWM-encoding hypothesis predicts higher probe recognition performance in the reported than in the non-reported and not shown conditions. In contrast, no such performance differences

are expected based on the type-activation hypothesis. According to this hypothesis, performance should be equal in the reported and non-reported conditions. More specifically, episodic short-term recognition should be possible for all presented memory letters, irrespective of their encoding into VWM. That is because all presented memory letters should have activated their visual types in visual long-term memory as part of the initial processing of the letters (e.g., Duncan and Humphreys, 1989; Bundesen, 1990; Bundesen et al., 2005; Kyllingsbæk, 2014; see above). Besides testing these hypotheses, Experiment 1 explored whether memory letters in the whole report task were encoded in a spatially clustered manner. That is, whether letters in close spatial proximity were encoded with preference over letters that were farther apart. Such a spatial clustering may reveal attentional selection strategies and this will become important in Experiment 2.

Method
Participants
Fourteen participants were paid to take part in the experiment. They were between 18 and 30 years old (*Mdn* = 20 years), nine were male, five female, 13 were right-handed and one left-handed, and all reported normal or corrected-to-normal visual acuity and color vision. All participants gave written informed consent before performing the experiments that were conducted according with the ethical standards of the German Psychological Association (Deutsche Gesellschaft für Psychologie, DGPs), and were approved by Bielefeld University's ethics committee. One additional participant was excluded from data analysis because of an experimentation error.

Apparatus and Stimuli
The experiment took place in a dimly lit room. Stimuli were presented on a 19″ CRT-screen (Trinitron MultiScan G420, Sony, Park Ridge, NJ, using a graphics card of type Quadro NVS 290, NVIDIA, Santa Clara, CA, USA) with a refresh rate of 85 Hz and a resolution of 1280 × 1024 pixels at physical dimensions of 36 cm × 27 cm. The participant's head was stabilized by a chin rest positioned 71.8 cm from the screen. Responses were collected using a standard computer keyboard with German layout. Labels indicating "yes" (by the German word "Ja") and "no" (by the German word "Nein") were placed above the F1 and F9 keys of the keyboard. The experiment was controlled by the Psychophysics Toolbox 3.0.12 extension (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) for MATLAB R2013b (The MathWorks, Natick, MA, USA).

A MAVOLUX-digital luminance meter (Gossen, Nuremberg, Germany) was used to measure stimulus luminance. Black letter stimuli ($0.32°$ of visual angle × $0.48°$; < 1 cd × m^{-2}) from the set [ABDEFGHJKLMNOPRSTVXZ] (this set of letters was chosen to avoid highly confusable letters, as e.g., by Poth et al., 2015) were located equally spaced on an imaginary circle with a radius of $2°$ around screen center. Fixation cross ($0.32° × 0.32°$) and response screen text were white (108 cd × m^{-2}). The response screen showed the German text "Buchstaben?", which

means "Letters?" in English. Stimuli were shown against a gray background (21 cd × m^{-2}).

Procedure and Design

Before the experiment, participants read instructions on the screen and reported them to the experimenter in their own words. The experimenter repeated the instructions again, if participants had reported them incorrectly. **Figure 1** illustrates the experimental paradigm. Participants initiated each trial by pressing the space-bar. In the beginning of a trial, a fixation cross was shown for 400 ms. Next, 10 memory letters were presented for 200 ms. The letters were randomly drawn without replacement from the set of used letters. The memory letters were followed by a blank interstimulus interval (ISI) lasting for 1000 ms (this duration ensures that early sensory (iconic) memory representations of the letters have been decayed, e.g., Sperling, 1960; Phillips, 1974; Irwin and Thomas, 2008), after which a response screen prompted participants to enter letters. Participants should report as many from the preceding memory letters as they could (without being required to report as many as 10 letters). A maximum of 10 letters could be entered (but this never happened). After confirming that they had finished reporting letters by pressing the enter-key, another ISI of 94 ms followed. Then a single probe letter was presented. Participants indicated whether or not this probe was one of the preceding memory letters by pressing the F1 or F9 key, respectively.

The probe was manipulated in three conditions of a within-subjects design. In the reported condition, the probe was randomly chosen from the letters which were shown and reported by the participant on this trial. In the non-reported condition, the probe was one of the letters that were shown on this trial but that the participant did not report. In both of these two conditions, probes appeared at their locations in the display of the memory letters. In the not shown condition, the probe was randomly chosen from the set of all letters excluding the memory letters of the trial (irrespective of whether participants had entered these letters). In this condition, the probe appeared at a random location.

Participants performed three blocks of 100 trials, each comprising 25 trials of the reported, 25 trials of the non-reported, and 50 trials of the not shown condition. Twice as many trials of the not shown as of the other two conditions were included to equate the number of trials in which a previously shown (correct answer "yes") or a not shown letter (correct answer "no") was probed. Within each block, trials of the three conditions were administered in random order. Participants performed twelve training trials prior to the experiment.

Results and Discussion

A significance criterion of $p < 0.05$ was used for all statistical analyses. Performance in the three conditions was compared using one-way repeated-measures analyses of variance with type II sums-of-squares for which η_G^2 (Bakeman, 2005) is reported as effect size. Where the assumption of sphericity was violated, p-values are based on Greenhouse–Geisser-corrected degrees of freedom and the correction factor ε is reported alongside

the uncorrected degrees of freedom. Paired t-tests (two-tailed) with Bonferroni-corrected p-values (p_B) were used for pairwise comparisons for which d_z (Cohen, 1988) is reported as effect size. These t-tests were supplemented with corresponding Bayes factors (BF; Rouder et al., 2009), of which values greater one favor the null hypothesis and values smaller one favor the alternative hypothesis. All analyses were performed using R (3.0.3; R Development Core Team, 2016).
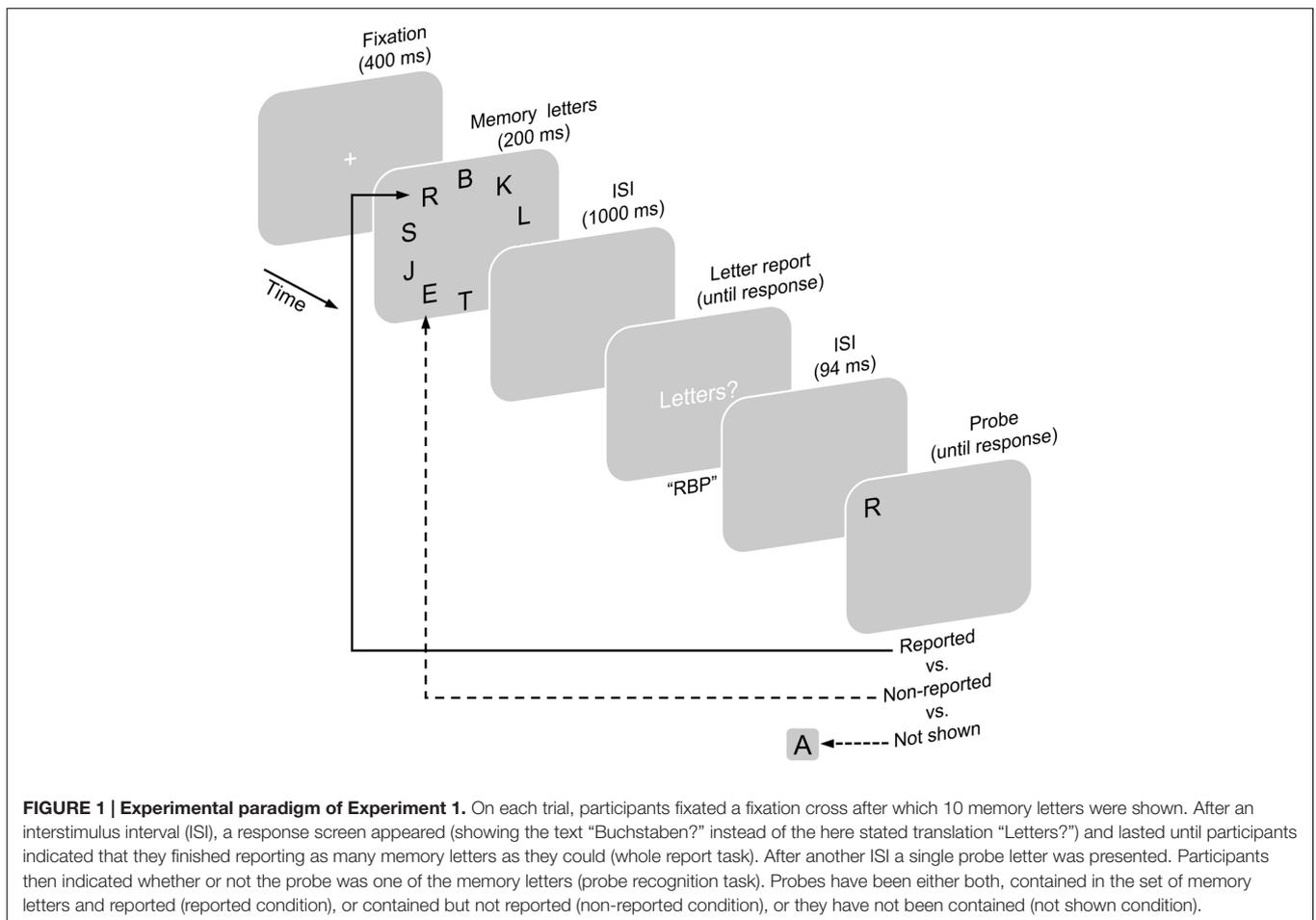
A total of 3.3% of all trials were discarded before analysis because either, (1) none of the memory letters was reported (0.57%), or (2) duplicate letters were contained in the letter report (2.76%).

Letter Report Performance

Letter report performance was assessed as participants' mean number of correctly reported letters, that is, for each individual participant the mean number of typed-in letters matching one of the memory letters across trials. There were no significant differences regarding letter report performance in the three conditions, $F(2,26) = 2.231$, $p = 0.128$, $\eta_G^2 = 0.002$. In addition, mean letter report performance was in the range of three to four letters in all three conditions (reported: $M = 3.62$, $SD = 0.59$, min = 2.41, max = 4.60; non-reported: $M = 3.56$, $SD = 0.61$, min = 2.35, max = 4.41; not shown: $M = 3.56$, $SD = 0.59$, min = 2.44, max = 4.5), consistent with previous estimates of VWM capacity in letter report tasks (Sperling, 1960; Shibuya and Bundesen, 1988).

Spatial Clustering of Reported Letters

Whether letters were encoded into VWM in a spatially clustered manner was assessed as follows. For each trial, the extent to which reported letters were spatially clustered within the original display of memory letters (i.e., their spatial proximity in this display) was quantified. The data was collapsed across conditions, since trials in the three conditions did not differ until after letters had been reported. Each correctly reported letter was selected for one step of the analysis. For this selected letter, it was determined whether or not the memory letters at the 10 positions relative to it were correctly reported (**Figure 2A**). This must be always the case for relative position zero, as this is the position of the selected letter itself. The procedure resulted in a matrix with the dimensions number of reported letters (rows) × 10 letter positions (columns) and with entries coding for whether or not a given letter has been reported. Now, spatial clustering of letter reports was assessed as the proportions of reported letters for each letter position (i.e., for each column) across all reported letters (i.e., across all rows). If participants reported letters in a spatially random manner, then these proportions should be equal with the exception of a proportion of 1 for the selected letters (see **Figure 2B** for a computer simulation). In contrast, spatial clustering in encoding letters would become manifest in higher proportions for letters at positions more proximal compared with positions more distant to the selected letter (**Figure 2C** for a computer simulation). Note that these analyses require that the number of presented letters clearly exceeds participants' VWM capacity because otherwise there would be no clear differences between proportions. This condition is assumed to be met because participants reported

FIGURE 1 | Experimental paradigm of Experiment 1. On each trial, participants fixated a fixation cross after which 10 memory letters were shown. After an interstimulus interval (ISI), a response screen appeared (showing the text "Buchstaben?" instead of the here stated translation "Letters?") and lasted until participants indicated that they finished reporting as many memory letters as they could (whole report task). After another ISI a single probe letter was presented. Participants then indicated whether or not the probe was one of the memory letters (probe recognition task). Probes have been either both, contained in the set of memory letters and reported (reported condition), or contained but not reported (non-reported condition), or they have not been contained (not shown condition).

between three and four of the 10 presented letters (see the letter report performance above).
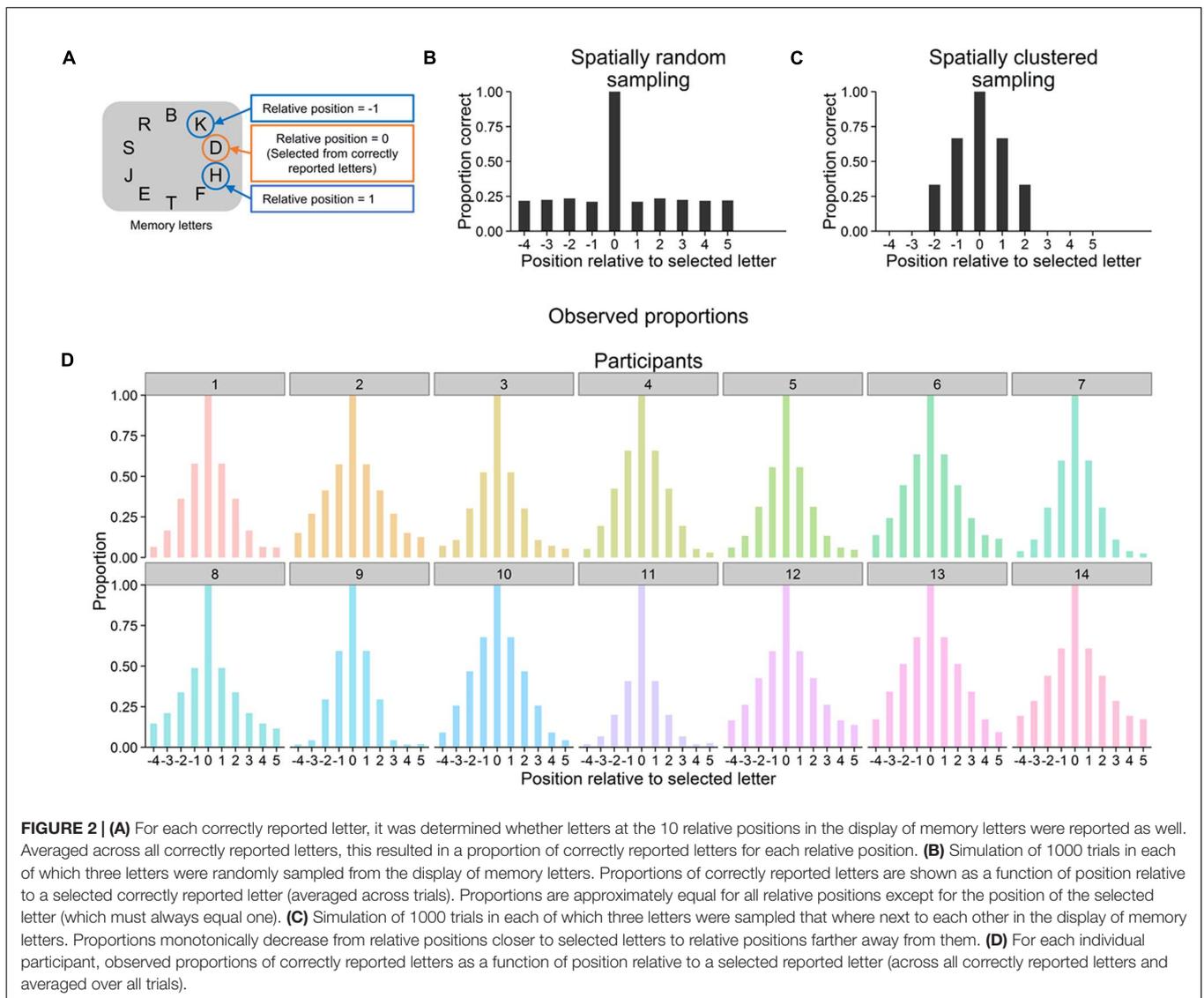
As can be seen in **Figure 2D**, the mean proportions of reported letters monotonically decreased with increasing distance to selected letters and this pattern was present in all participants. Page's trend test was used to test whether monotonic decreases from closer to more distant positions were statistically significant. To this end, Page's trend test was applied to the participants' proportions at relative positions -1 to -4 and, separately, at relative positions 1 to 5 (**Figure 2A**). Results revealed monotonic decreases for both of these subsets of the data, locations -1 to -4: $L = 420$, $p < 0.001$, locations 1 to 5: $L = 768$, $p < 0.001$ (and these monotonic decreases were present in all of the three blocks of trials, all $Ls > = 420$, all $ps < 0.001$).

Selective encoding of letters into VWM was not spatially random. Instead, all participants encoded subsets of the memory letters into VWM that were in close spatial proximity in the letter display. This spatial clustering may reflect an attentional encoding strategy. Participants learned over trials that always more memory letters were shown than they could report. Thus, participants learned they had to select subsets of the memory letters for report. Spatial clustering may be a means to accomplish such a selection from equally task-relevant objects by restricting encoding to objects in close spatial proximity. In this way, spatial

clustering may reflect the distribution of spatial attention (e.g., Posner, 1980; Bundesen, 1990), which in this specific case selects objects at or close to a strategically and internally specified location.

Probe Recognition Performance

Probe recognition performance was assessed as the proportion of trials on which probe letters were correctly recognized as having been shown or not shown on the trial. **Figure 3** depicts the participants' probe recognition performance, both at the sample and individual level. Probe recognition performance differed significantly between the three conditions, $F(2,26) = 44.912$, $\varepsilon = 0.522$, $p < 0.001$, $\eta_G^2 = 0.771$. Probe recognition performance was significantly higher in the reported ($M = 0.96$, $SD = 0.03$) compared with the non-reported ($M = 0.29$, $SD = 0.19$), $t(13) = 12.774$, $p_B < 0.001$, $d_z = 3.41$, $BF = 8.8 \times 10^{-7}$, and the not shown condition ($M = 0.74$, $SD = 0.20$), $t(13) = 4.170$, $p_B = 0.003$, $d_z = 1.11$, $BF = 0.028$. Moreover, performance was significantly lower in the non-reported than in the not shown condition, $t(13) = -4.498$, $p_B = 0.002$, $d_z = -1.20$, $BF = 0.016$. One-sample t-tests (two-sided) revealed that performance was significantly below the chance level of 0.5 in the non-reported condition, $t(13) = -4.243$, $p < 0.001$, $BF = 0.025$, whereas it was significantly above
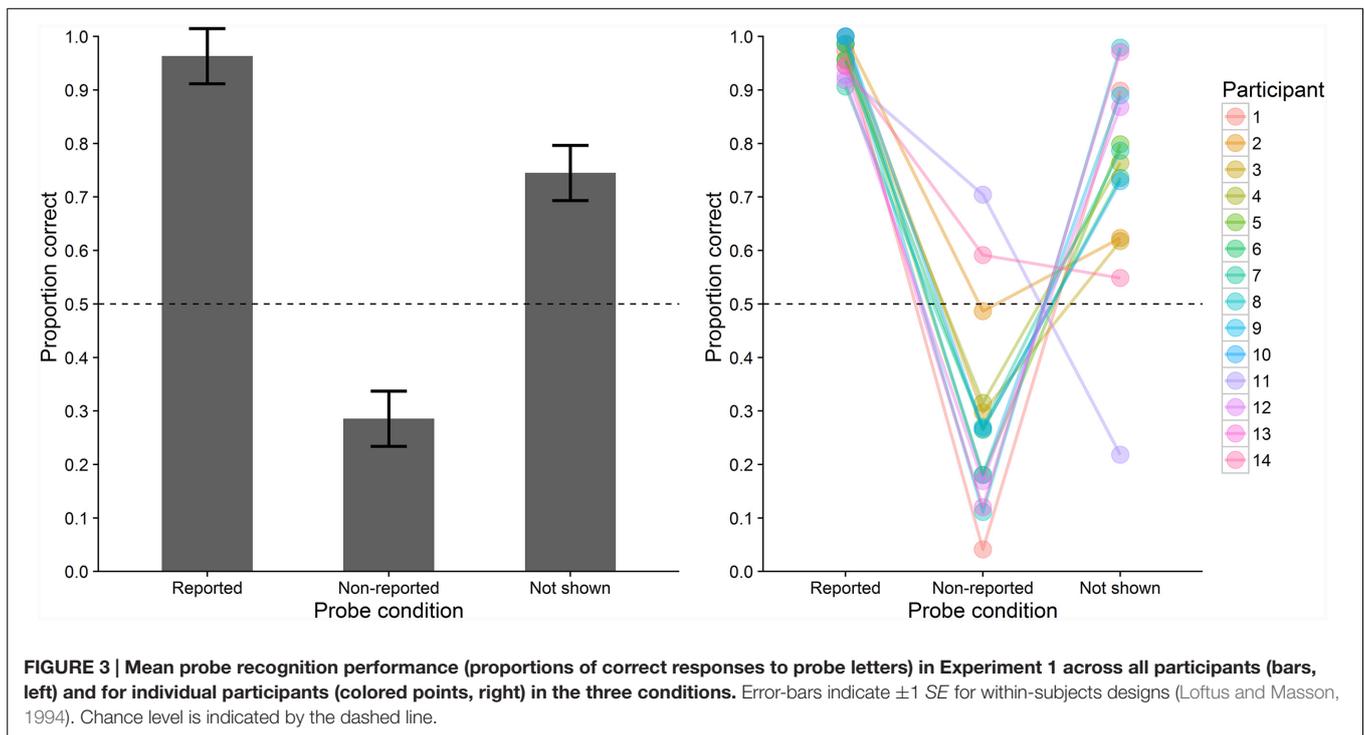
FIGURE 2 | (A) For each correctly reported letter, it was determined whether letters at the 10 relative positions in the display of memory letters were reported as well. Averaged across all correctly reported letters, this resulted in a proportion of correctly reported letters for each relative position. **(B)** Simulation of 1000 trials in each of which three letters were randomly sampled from the display of memory letters. Proportions of correctly reported letters are shown as a function of position relative to a selected correctly reported letter (averaged across trials). Proportions are approximately equal for all relative positions except for the position of the selected letter (which must always equal one). **(C)** Simulation of 1000 trials in each of which three letters were sampled that where next to each other in the display of memory letters. Proportions monotonically decrease from relative positions closer to selected letters to relative positions farther away from them. **(D)** For each individual participant, observed proportions of correctly reported letters as a function of position relative to a selected reported letter (across all correctly reported letters and averaged over all trials).

chance in the not shown condition, $t(13) = 4.589$, $p < 0.001$, $BF = 0.014$.

Whether probe recognition depended on how many letters participants entered for the whole report (irrespective of whether letters were correct) was assessed as the point-biserial correlation between the number of entered letters and probe recognition performance, separately for each participant and each condition. Values of three participants in the reported condition had to be excluded from this analysis because probe recognition was correct in all trials so that no correlation could be computed. One-sample t-tests (two-sided) indicated that the correlations of the 11 remaining participants did not significantly depart from zero in any of the three conditions, all $|ts|$ (10) < 1.713, all $ps > 0.110$, all $BFs > 1.149$.

Probe recognition performance was close to ceiling in the reported condition but it was substantially lower in the non-reported and not shown conditions. These findings clearly argue against the type-activation hypothesis which predicts

equal performance for all presented memory letters and hence equal performance in the reported and non-reported condition. Instead, the findings seem to support the VWM-encoding hypothesis which predicts higher performance in the reported condition, in which probe letters were encoded into VWM. However, before arriving at these conclusions, several issues should be considered. According to the VWM-encoding hypothesis, performance should have been at chance level in the non-reported condition but it was below chance level. This may indicate that participants based their probe responses not only on the letters they remembered having viewed on this trial. Rather, they may have partly based their responses on the letters they remembered having reported on this trial. This would have biased them away from responding those probes had been contained in the memory letters when they had not reported the letters of these probes. This bias might also have contributed to the above-chance performance in the not shown condition. Besides biasing responses, reporting the letters itself

FIGURE 3 | Mean probe recognition performance (proportions of correct responses to probe letters) in Experiment 1 across all participants (bars, left) and for individual participants (colored points, right) in the three conditions. Error-bars indicate ±1 *SE* for within-subjects designs (Loftus and Masson, 1994). Chance level is indicated by the dashed line.

might also have improved their subsequent episodic short-term recognition compared to non-reported letters. Similarly, reporting memory letters might have interfered with retaining non-reported letters. In addition, reporting the letters may have prolonged the interval that the non-reported letters had to be retained. In all of these cases, letters that were inaccessible for report might have been available for later episodic short-term recognition if intervening report requirements were controlled for. Therefore, the aim of Experiment 2 was to control for all effects reporting letters might have on probe recognition performance.

EXPERIMENT 2

Experiment 2 was designed to investigate episodic short-term recognition performance for letters that were more likely to be encoded into VWM compared with letters whose encoding was less likely. To manipulate the likelihood of encoding specific letters into VWM, we made use of the spatial clustering of VWM encoding found in Experiment 1. Participants briefly viewed a display of 10 letters in which a colored frame identified one letter as report-target and frames in a different color identified the nine other letters as non-targets *regarding report*. Participants' task was to report the single report-target after a retention interval. After reporting, a single probe letter was shown and participants were to indicate whether or not it had been presented as one of the preceding letters (**Figure 4**). There were three conditions. In the *report-target condition*, the probe tested recognition of the report-target. In the *near non-target condition*, the probe tested recognition of a letter that has been located directly beside the
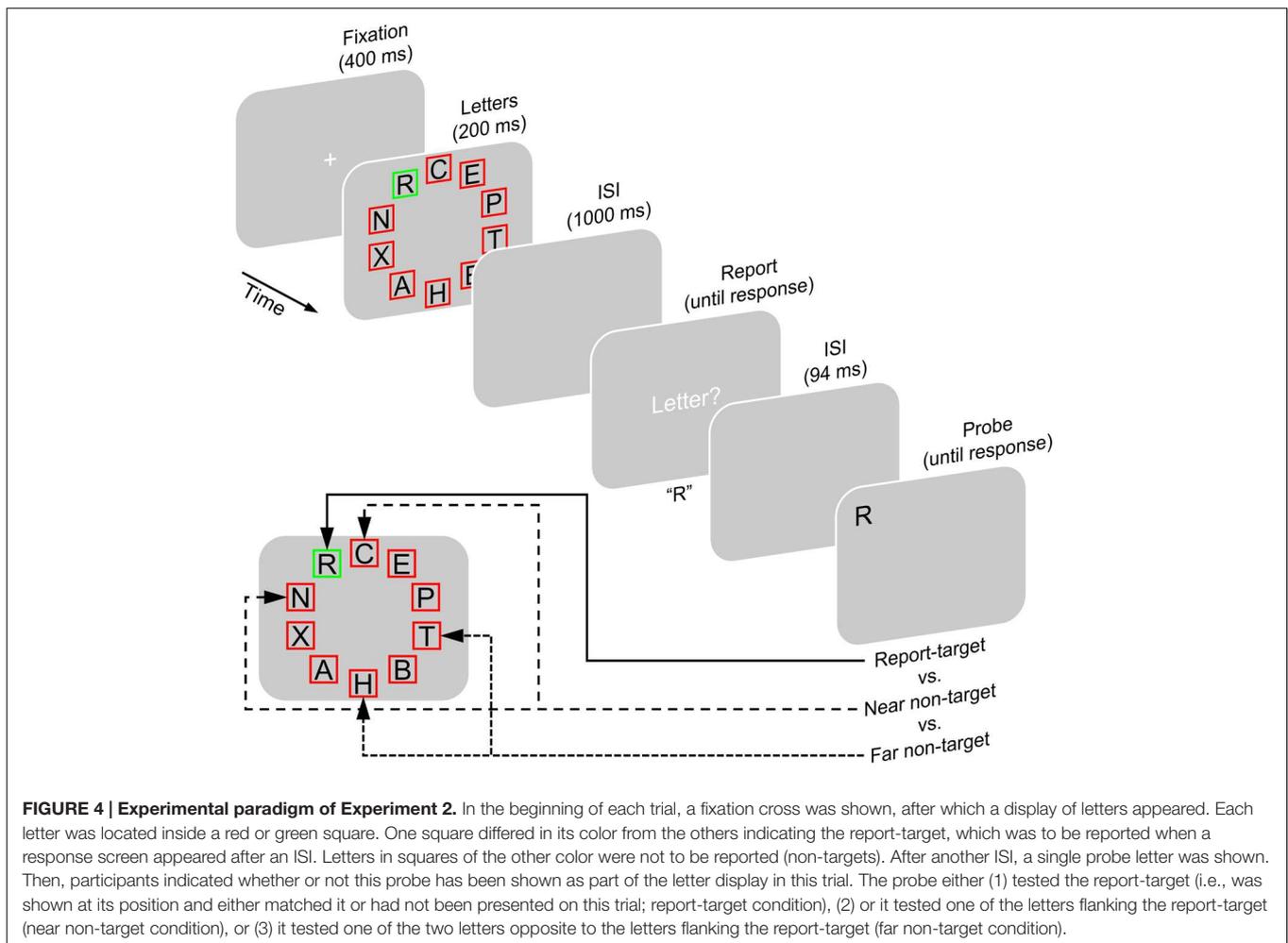
report-target. In the *far non-target condition*, the probe tested recognition of a letter that has been located far away from the report-target, on the other side of the letter display.

The report-target has to be encoded into VWM, in order to be accessible for being reported (e.g., Bundesen, 1990; Bundesen et al., 2005; Schneider, 2013). Because of the spatial clustering of letter reports in Experiment 1, we assumed that while participants aimed at encoding the report-target, they were more likely to encode near non-targets selectively compared with far non-targets. This is compatible with the view that spatial attention was primarily directed at the report-target (e.g., Kim and Cave, 1995; Gaspelin et al., 2015), but was secondarily directed more at near non-targets than at far non-targets or was secondarily directed at near non-targets only. According to the VWM-encoding hypothesis, probe recognition performance should be highest for report-targets, followed by near non-targets, and lowest for far non-targets because of their lowest likelihood of being encoded into VWM. In contrast, according to the type-activation hypothesis probe recognition performance should be equal for all presented letters and thus equal in all three conditions. Importantly, the near and far non-targets were not subject to report requirements.

Method
Participants

Ten paid participants took part in Experiment 2. They were between 22 and 30 years old (*Mdn* = 25). Four of them were male, six female, nine were right, and one left-handed. All participants reported normal or corrected-to-normal visual acuity and color vision. They gave written informed consent before performing the experiments that were conducted according to the ethical

FIGURE 4 | Experimental paradigm of Experiment 2. In the beginning of each trial, a fixation cross was shown, after which a display of letters appeared. Each letter was located inside a red or green square. One square differed in its color from the others indicating the report-target, which was to be reported when a response screen appeared after an ISI. Letters in squares of the other color were not to be reported (non-targets). After another ISI, a single probe letter was shown. Then, participants indicated whether or not this probe has been shown as part of the letter display in this trial. The probe either (1) tested the report-target (i.e., was shown at its position and either matched it or had not been presented on this trial; report-target condition), (2) or it tested one of the letters flanking the report-target (near non-target condition), or (3) it tested one of the two letters opposite to the letters flanking the report-target (far non-target condition).

standards of the German Psychological Association (DGPs), and were approved by Bielefeld University's ethics committee.

Apparatus and Stimuli

The apparatus and experimental setup of Experiment 2 were the same as those of Experiment 1. The stimuli of Experiment 2 were identical to those of Experiment 1 with the following exceptions. All letters were placed inside a square frame ($0.72° \times 0.72°$). Frames of the nine non-targets were either all red (20 cd $\times$ m^{-2}; RGB: 255, 0, 0) or green (76 cd $\times$ m^{-2}; RGB: 0, 255, 0). The frame of the report-target was in the other color (i.e., green when the others were red or red when the others were green). The colors of report-target and non-targets remained the same throughout the experiment. Whether red or green indicated the report-target was counterbalanced across the sample. The text of the response screen was identical to that in Experiment 1, except that it prompted participants to enter only one instead of several letters (by the German text "Buchstabe?", which means "Letter?" in English).

Procedure and Design

As illustrated in **Figure 4**, the experimental paradigm of Experiment 2 was identical to that of Experiment 1 except for

the following aspects. Instead of all 10 letters, participants were to report only the one report-target (partial report). On each trial, the position of the report-target was randomly chosen. No confirmation of this report was required, instead the trial proceeded as soon as a letter-key had been pressed. As in Experiment 1, at the end of each trial a single probe letter was shown and participants were required to indicate whether or not it was shown within the letter display of this trial. Participants performed three conditions of a within-subjects design. In the *report-target condition*, the probe appeared at the location of the report-target and either matched the report-target or consisted in a letter not presented on this trial. In the *near non-target condition*, the probe appeared at the location of one of the two letters that flanked the report-target and either matched this letter or had not been presented on this trial. In the *far non-target condition*, the probe appeared at the location of one of the two letters opposite to the two flanking letters, on the other side of the letter display than the report-target and either matched this letter or had not been shown on this trial.

Participants performed four blocks of 72 trials each comprising 24 trials of the report-target, near non-target, and far non-target condition. For the two non-target conditions,

probes appeared equally often at positions in clockwise or counter-clockwise direction of the report-target. In each of the three conditions and for each possible probe location, trials with probes matching the former letter at the probe's location (correct answer "yes") and probes not shown (correct answer "no") occurred equally often. Participants performed 24 training trials prior to the experiment.

Results and Discussion

The same statistical procedures were used as in Experiment 1. Two trials were excluded from analysis because participants entered more than one letter in their letter report (which could happen only if participants pressed two keys close to simultaneously). Whether report-targets were in red or green frames did not interact with any of the below described dependent variables, all Fs < 1.64, all ps > 0.227 (revealed by a repeated-measures ANOVA with type III sums-of-squares). Therefore, data of participants with report-targets in red and green frames was collapsed for the following analyses.

Letter Report Performance

Letter report performance was assessed as participants' proportion of trials on which the report-target was correctly reported. Unsurprisingly, there were no significant differences between letter report performance in the three experimental probe conditions (report-target condition: $M = 0.94$, $SD = 0.07$; near non-target condition: $M = 0.94$, $SD = 0.06$; far non-target condition: $M = 0.93$, $SD = 0.07$), $F(2,18) = 0.545$, $p = 0.589$, $\eta_G^2 < 0.004$. In addition, Friedman's test was applied, because the assumption of normal distribution of the repeated-measures analysis of variance was not met. This test yielded a non-significant effect as well, $\chi^2(2) = 1.316$, $p = 0.518$.

Participants' letter report performance did not differ reliably between the three conditions. Participants achieved close-to-ceiling performance in all three conditions, as could be expected since only one letter had to be reported which should not touch the capacity limit of VWM (Sperling, 1960; Shibuya and Bundesen, 1988).

Probe Recognition Performance

Different from Experiment 1, each condition contained trials in which probes did and trials in which probes did not match the letters they referred to. Therefore, probe recognition performance could be quantified as d', the difference between the z-transformed rate of correct responses to probes shown on this trial, z("hit rate"), and the z-transformed rate of false responses

to probes not shown on this trial, z("false alarm rate"; for an overview, see Macmillan and Creelman, 2005). Performance at chance level leads to a d' of zero and close to perfect performance to values of 4.65 (or higher and 0.5 was added to all data cells on which hit and false alarm rates were based to avoid infinite values for d', Macmillan and Creelman, 2005, pp. 8–9,). To facilitate comparison with the results of Experiment 1, in **Table 1** we also report the probe recognition performance assessed as the proportion of trials on which probe letters were correctly recognized as having been shown or not shown on the trial.
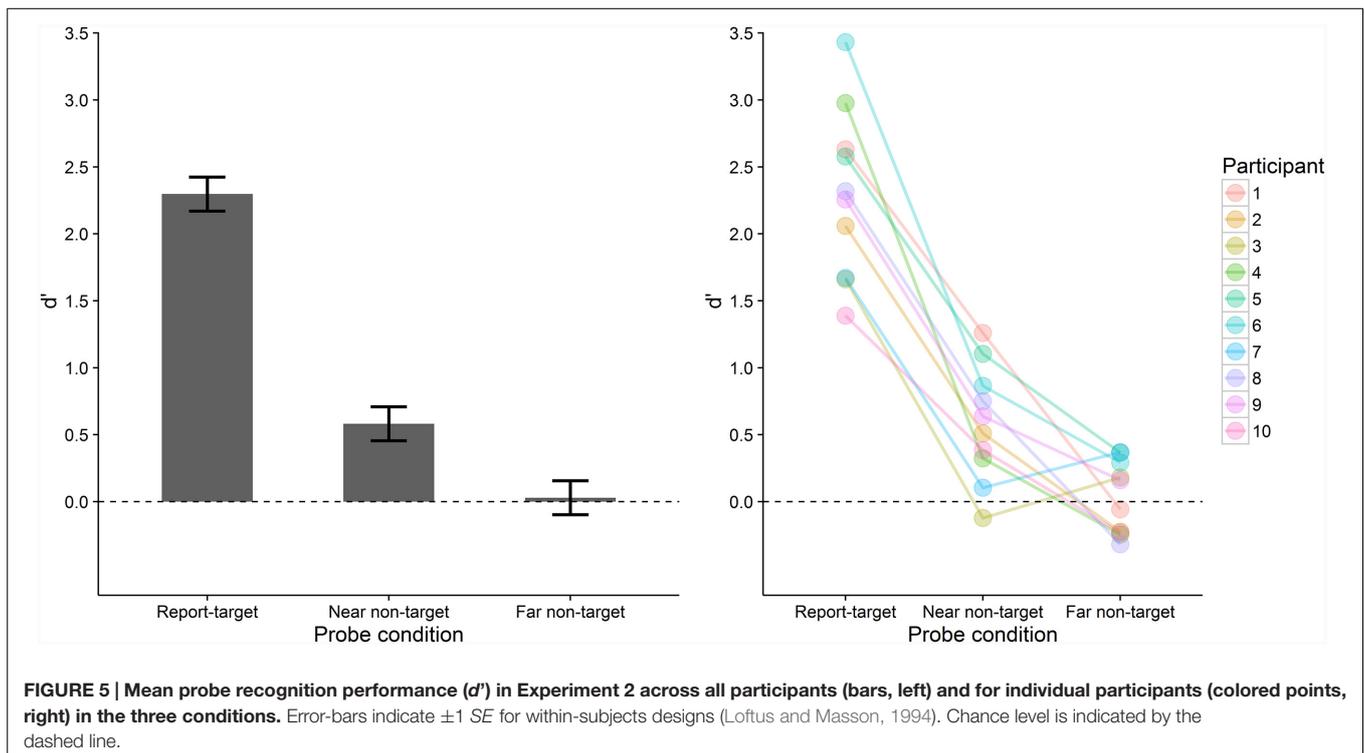
Figure 5 depicts participants' probe recognition performance in the three conditions at the sample and individual level. Performance differed significantly between the three conditions, $F(2, 18) = 86.859$, $p < .001$, $\eta_G^2 = 0.824$. That is, performance was significantly higher in the report-target ($M = 2.30$, $SD = 0.63$) compared with the near non-target ($M = 0.58$, $SD = 0.43$), $t(9) = 10.562$, $p_B < 0.001$, $d_z = 3.34$, $BF = 1.3 \times 10^{-4}$, and far non-target condition ($M = 0.03$, $SD = 0.27$), $t(9) = 10.770$ $p_B < 0.001$, $d_z = 3.41$, $BF = 1.2 \times 10^{-4}$. Performance was also significantly higher in the near than in the far non-target condition, $t(9) = 3.435$, $p_B = 0.022$, $d_z = 1.09$, $BF = 0.127$. This data pattern was present in all except two participants whose performance was slightly higher in the far compared with the near non-target condition (**Figure 5**, right). In addition, two one-sample t-tests revealed that performance was significantly above chance in the near non-target condition, $t(9) = 4.262$, $p = 0.002$, $BF = 0.045$, but did not differ from chance level (i.e., a d' of zero) in the far non-target condition, $t(9) = 0.335$, $p > 0.745$, $BF = 3.086$.

Probe recognition performance was highest when the probe letter tested the former report-target which had been encoded into VWM, as evident from the near-ceiling performance in reporting its identity. Importantly, performance was higher for near non-targets than for far non-targets. This indicates that episodic short-term recognition was better for letters that were more likely to be encoded into VWM compared with letters less likely to be encoded (given that encoding into VWM seems to proceed in a spatially clustered manner, see Experiment 1). In fact, performance for far non-targets was at chance level which suggests that episodic short-term recognition was not possible for these letters. Furthermore, Experiment 2 controlled for potential alternative explanations of the findings of Experiment 1. These alternative explanations stated that differences between the conditions did not stem from whether letters were encoded into VWM but from whether letters were reported. In Experiment

TABLE 1 | Probe recognition performance in the three conditions of Experiment 2 assessed as the proportion of correct responses to the probe.

	M (SD)	vs. report-target	vs. near non-target
Report-target	0.79 (0.09)	–	–
Near non-target	0.61 (0.08)	$t(9) = -6.515$, $p_B < 0.001$, $d_z = -2.06$, $BF = 3.79 \times 10^{-3}$	–
Far non-target	0.51 (0.05)	$t(9) = -9.313$, $p_B < 0.001$, $d_z = -2.94$, $BF = 3.30 \times 10^{-4}$	$t(9) = -3.451$, $p_B = 0.022$, $d_z = -1.09$, $BF = 0.125$

Provided are means (M) with standard deviations (SD) in parentheses, and for each of the three conditions (rows), the results of the pairwise comparisons to the other two conditions (last two columns).

FIGURE 5 | Mean probe recognition performance (*d'*) in Experiment 2 across all participants (bars, left) and for individual participants (colored points, right) in the three conditions. Error-bars indicate ±1 *SE* for within-subjects designs (Loftus and Masson, 1994). Chance level is indicated by the dashed line.

2, near and far non-targets both did not have to be reported and differed only in their distance from the report-target. Hence, the performance difference between these two conditions cannot be attributed to effects reporting letters itself might have on performance. Therefore, we interpret the higher performance for near non-targets compared with the performance at chance level for far non-targets as strong evidence for the VWM-encoding hypothesis. Conversely, we interpret this finding as evidence against the type-activation hypothesis.

GENERAL DISCUSSION

We investigated whether episodic short-term recognition of objects from a previous processing episode requires that these objects have been encoded into VWM. For this purpose, we introduced a new paradigm combining letter report with probe recognition. In two experiments, episodic short-term recognition was assessed as performance in recognizing whether a probe letter was presented in the preceding letter display of the current trial. In Experiment 1, probe recognition performance was higher for letters that had been encoded into VWM compared with letters that had not been encoded. In Experiment 2, only a single letter had to be reported in the letter report task. This controlled for effects reporting letters itself might have on probe recognition. In Experiment 2, probe recognition performance was higher for non-target letters that were near to a report-target letter, and hence more likely to be encoded into VWM, compared with non-target letters far from the report-target, whose encoding was less likely. Crucially, this difference in probe recognition refers to non-target letters which did not have to be reported. Strikingly,

performance was at chance level for letters far from the report-target which were unlikely to enter VWM. Therefore, we interpret the present findings as strong evidence for the VWM-encoding hypothesis which states that episodic short-term recognition presupposes that visual objects have been encoded into VWM. Conversely, we interpret these findings as evidence against the type-activation hypothesis. Note that one might distinguish a strong form of the type-activation hypothesis, the one that we have put forward so far, and a weaker form. The strong form states that episodic short-term recognition can be accomplished perfectly (at least in principle) for all objects of the current visual field. In contrast, the weaker form states that episodic short-term recognition can be accomplished for all objects of the visual field, but not perfectly, and that recognition performance may be improved by additional encoding into VWM. The results of Experiment 2 provide evidence against both forms of the type-activation hypothesis. The finding that probe recognition performance was higher for near than for far non-targets argues against the strong form. The finding that performance was at chance level in the far non-target condition argues against the weak form. That is, episodic short-term recognition seemed impossible in this condition. Thus, taken together, the present findings indicate that type-activation is not sufficient for later episodic short-term recognition but that encoding into VWM is required instead.

Visual Working Memory as a Basis of Episodic Short-Term Recognition

Encoding an object into VWM seems to be necessary for its later episodic short-term recognition. This means that the functional

basis of episodic short-term recognition emerges at a level of processing after the activation of visual types in visual long-term memory (e.g., Kanwisher, 1987; Kahneman et al., 1992; cf. Schneider, 1995) and after visual attention has mediated selective encoding into VWM (Duncan and Humphreys, 1989; Bundesen, 1990; Bundesen et al., 2005; Schneider, 2013). In the present study, letters were used as visual objects. After successful visually based recognition, letters can be processed verbally, which makes it likely that their episodic short-term recognition also involved verbal processing in addition to visual processing. However, because the letters had to be acquired visually, they had to be encoded into VWM first, before such a verbal processing could take place. After their encoding into VWM, they may have been recoded into a verbal format. Such a verbal format may have provided the advantage of verbal rehearsal by verbal working memory, which may have prolonged and secured their retention (e.g., Logie, 2011; Baddeley, 2012). Thus, importantly, even though episodic short-term recognition may rely on several different (working) memory mechanisms (such as visual and verbal ones), encoding into VWM seems to be a necessary processing step for these mechanisms to operate.

Why may encoding into VWM be necessary for episodic short-term recognition? Several theories assume that by encoding into VWM, information about visual objects is transformed into a special representational state (e.g., Cowan, 1988; Oberauer, 2002; LaRocque et al., 2014; cf. Olivers et al., 2011). We suggest that it is this representational state that makes encoding into VWM a requirement of episodic short-term recognition. Specifically, we propose that two characteristics of this representational state are necessary for episodic short-term recognition: *binding* and *robustness*.

Binding means that different visual features of an object are integrated which yields representations of objects as a whole, with all their features (e.g., Treisman and Gelade, 1980). The mere presentation of objects activates visual types (features) in visual long-term memory but this happens in isolation (cf. Bundesen, 1990; Schneider, 1995). Episodic short-term recognition requires binding of activated visual types because otherwise objects that share visual features cannot be distinguished. VWM is assumed to mark the first level in the course of visual processing at which the visual types (or features) activated by an object are bound to integrated object representations (Bundesen, 1990; Luck, 2008; Schneider, 2013; Kyllingsbæk, 2014). This point is illustrated by referring to integrated object representations as VWM objects (Schneider, 2013), which have also been called object files (Kahneman et al., 1992) and visual tokens (Schneider, 1995). In sum, the binding of visual types within object representations in VWM may be one reason for that episodic short-term recognition requires encoding into VWM.

Robustness means that object representations in VWM are protected against so-called proactive interference (Keppel and Underwood, 1962). Proactive interference arises when the same visual objects occur repeatedly (e.g., Endress and Potter, 2014). It describes an impairment in recognizing if an object has been viewed in the very recent past as opposed to having been encountered before at all (e.g., Endress and Potter, 2014). Episodic short-term recognition clearly requires

to assess whether an object has been viewed in a recently passed episode rather than at some unspecified point in the past. Hence, successful episodic short-term recognition presupposes that proactive interference is eliminated. Robustness against proactive interference is assumed to be a hallmark of VWM representations and providing it is considered a core function of VWM (Endress and Potter, 2014). Thus, taken together, episodic short-term recognition may presuppose encoding of objects into VWM because this might establish representations of objects as bound units (cf. Luck and Vogel, 1997) which are robust against proactive interference (cf. Endress and Potter, 2014).

Episodic Short-Term Recognition Might Be Constrained by an Encoding-Limitation but Not a Retention-Limitation of Visual Working Memory

As we have argued, the present findings indicate that episodic short-term recognition presupposes encoding into VWM but this seems to conflict with earlier findings. Specifically, Sternberg (1966) presented participants with series of up to six digits followed by a probe digit. Participants indicated whether the probe was contained in a given series. The six presented digits exceed the number of about three to four objects that VWM can hold (e.g., Sperling, 1960; Shibuya and Bundesen, 1988; Luck and Vogel, 1997). Thus, when the last two digits were shown, VWM should have already been filled up so that the digits could not be encoded into VWM. Nevertheless, Sternberg found that probe recognition performance was close to ceiling even for six digits. One might attribute this result to the relatively long presentation durations of digits (1.2 s) that could have allowed verbal rehearsal (e.g., Sternberg, 1975). However, congruent to Sternberg's findings, later experiments revealed high levels of probe recognition performance for objects that were presented more briefly and thus difficult to rehearse verbally (Endress and Potter, 2014). Taken together, these findings are compatible with the type-activation hypothesis in that they suggest episodic short-term recognition is possible also for objects that have not reached VWM.

How may the conflict between the present and Sternberg's (1966; cf. Endress and Potter, 2014) findings be resolved? One solution is provided by Schneider's (2013) recent "theory of task-driven visual attention and working memory" (TRAM) which offers an account of how visual information processing might be accomplished within and across processing episodes. According to TRAM, a new processing episode is started with each onset of visual objects (e.g., after a saccadic eye movement). A processing episode comprises three phases. Premising upon Bundesen's (1990) theory of visual attention (a model of biased competition, Desimone and Duncan, 1995), TRAM's first two phases describe how visual attention mediates selective encoding of visual objects into capacity-limited VWM. In TRAM's third phase, objects that have been encoded into initial activation-based VWM (i.e., VWM based on persistent neural activity) are consolidated which results in *passive* VWM representations

(which do not require neural activity but may rely on short-term changes in synaptic connectivity, as reviewed by Eriksson et al., 2015; Postle, 2015; and Stokes, 2015). Critically, according to TRAM, the number of passive VWM representations is not constrained by the traditionally assumed capacity-limitation of VWM. With this in mind, one may interpret classical estimates of VWM capacity (Sperling, 1960; Shibuya and Bundesen, 1988; Luck and Vogel, 1997) as reflecting an encoding limitation but not a retention limitation. In other words, classical VWM capacity may constrain the amount of object information that can be acquired within one processing episode but not the amount of information that can be retained across episodes. In Sternberg's (1966) paradigm, each of the serially presented digits should have started a new processing episode. Within each of these episodes, a passive VWM representation of the digit should have emerged. Probe recognition should then have been based on a comparison of these passive VWM representations with actual probe digits (which could involve retrieving passive representations again into classical activation-based VWM; Schneider, 2013). In this vein, episodic short-term recognition becomes possible for more serially presented objects than classical VWM can retain. In contrast, TRAM posits that if several objects are presented simultaneously, as in the present experiments, then this can reach the encoding limit of VWM. All simultaneously presented objects are processed within the same processing episode. Therefore, encoding further objects becomes impossible if activation-based VWM is filled up. Critically, creating passive VWM representations of objects presupposes that the objects have been encoded into VWM. Thus, in a given processing episode, only as many objects as VWM can hold can be consolidated into passive VWM representations. As a consequence, episodic short-term recognition across successive processing episodes should be limited with respect to the number of simultaneously shown objects that can be encoded into VWM. In contrast, episodic short-term recognition should not be restricted with respect to the number of retained objects in VWM because this includes also passive VWM representations that have arisen over the course of several episodes, as in Sternberg's experiments. Interestingly, recent findings might suggest that in such situations of serial object presentations (RSVP), the capacity of passive VWM can be extended beyond "magical number four" by eliminating proactive interference (Endress and Potter, 2014). As an alternative to consolidation in passive VWM, representations of objects in classical VWM could also be recoded into a different representational format (Petersen et al., 2012) which might

then be used for later episodic short-term recognition. The objects of the present experiments consisted of letters which may have been recoded into the verbal format (that is open to verbal rehearsal, e.g., Sternberg, 1975, and may allow retention by working memory systems dedicated to verbal information, e.g., Baddeley, 2012). However, since the to-be-recoded object information is acquired visually, recoding would still presuppose encoding into VWM (Petersen et al., 2012). Hence, episodic short-term recognition would still be constrained by the encoding limitation of VWM but not by a retention limit. However, testing this hypothesis is left for further experimental studies.

CONCLUSION

The present study shows that episodic short-term recognition of objects from previous episodes presupposes that the objects have been processed up to the level of VWM. In this way, VWM not only provides bound visual objects for online perception and action within a processing episode but also paves the way for episodic short-term recognition across episodes. However, this also implies that episodic short-term recognition is only possible for a limited number of simultaneously presented objects due to the encoding limitation of VWM (Schneider, 2013; cf. Sperling, 1960; Shibuya and Bundesen, 1988; Luck and Vogel, 1997).

AUTHOR CONTRIBUTIONS

CP and WS designed the research. CP programmed the experiments and analyzed the data. CP and WS wrote the paper.

FUNDING

ACKNOWLEDGMENT

REFERENCES

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29. doi: 10.1146/annurev-psych-120710-100422

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behav. Res. Methods* 37, 379–384. doi: 10.3758/BF03192707

Bays, P. M., and Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science* 321, 851–854. doi: 10.1126/science.1158023

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357

Bundesen, C. (1990). A theory of visual attention. *Psychol. Rev.* 97, 523–547. doi: 10.1037/0033-295X.97.4.523

Bundesen, C., and Habekost, T. (2008). *Principles of Visual Attention: Linking Mind and Brain.* Oxford: Oxford University Press.

Bundesen, C., Habekost, T., and Kyllingsbæk, S. (2005). A neural theory of visual attention: bridging cognition and neurophysiology. *Psychol. Rev.* 112, 291–328. doi: 10.1037/0033-295X.112.2.291

Bundesen, C., Habekost, T., and Kyllingsbæk, S. (2011). A neural theory of visual attention and short-term memory (NTVA). *Neuropsychologia* 49, 1446–1457. doi: 10.1016/j.neuropsychologia.2010.12.006

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. New York, NY: Psychology Press.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychol. Bull.* 104, 163–191. doi: 10.1037/0033-2909.104.2.163

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–185. doi: 10.1017/S0140525X01003922

Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222. doi: 10.1146/annurev.ne.18.030195.001205

Donkin, C., and Nosofsky, R. M. (2012). The structure of short-term memory scanning: an investigation using response time distribution models. *Psychon. Bull. Rev.* 19, 363–394. doi: 10.3758/s13423-012-0236-8

Duncan, J. (2006). EPS Mid-Career Award 2004: brain mechanisms of attention. *Q. J. Exp. Psychol.* 59, 2–27. doi: 10.1080/17470210500260674

Duncan, J. (2013). The structure of cognition: attentional episodes in mind and brain. *Neuron* 80, 35–50. doi: 10.1016/j.neuron.2013.09.015

Duncan, J., and Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychol. Rev.* 96, 433–458. doi: 10.1037/0033-295X.96.3.433

Dyrholm, M., Kyllingsbæk, S., Espeseth, T., and Bundesen, C. (2011). Generalizing parametric models by introducing trial-by-trial parameter variability: the case of TVA. *J. Math. Psychol.* 55, 416–429. doi: 10.1016/j.jmp.2011.08.005

Endress, A. D., and Potter, M. C. (2014). Large capacity temporary visual memory. *J. Exp. Psychol. Gen.* 143, 548–565. doi: 10.1037/a0033934

Eriksson, J., Vogel, E. K., Lansner, A., Bergström, F., and Nyberg, L. (2015). Neurocognitive architecture of working memory. *Neuron* 88, 33–45. doi: 10.1016/j.neuron.2015.09.020

Gaspelin, N., Leonard, C. J., and Luck, S. J. (2015). Direct evidence for active suppression of salient-but-irrelevant sensory inputs. *Psychol. Sci.* 26, 1740–1750. doi: 10.1177/0956797615597913

Henderson, J. M. (1994). Two representational systems in dynamic visual identification. *J. Exp. Psychol. Gen.* 123, 410–426. doi: 10.1037/0096-3445.123.4.410

Henderson, J. M., and Anes, M. D. (1994). Roles of object-file review and type priming in visual identification within and across eye fixations. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 826–839. doi: 10.1037/0096-1523.20.4.826

Herwig, A. (2015). Linking perception and action by structure or process? Toward an integrative perspective. *Neurosci. Biobehav. Rev.* 52, 105–116. doi: 10.1016/j.neubiorev.2015.02.013

Irwin, D. E., and Thomas, L. E. (2008). "Visual sensory memory," in *Visual Memory*, eds S. J. Luck and A. Hollingworth (Oxford: Oxford University Press), 9–41.

Jensen, O., and Lisman, J. E. (1998). An oscillatory short-term memory buffer model can account for data on the Sternberg task. *J. Neurosci.* 18, 10688–10699.

Kahana, M. J., and Sekuler, R. (2002). Recognizing spatial patterns: a noisy exemplar approach. *Vision Res.* 42, 2177–2192. doi: 10.1016/S0042-6989(02)00118-9

Kahneman, D., Treisman, A., and Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cogn. Psychol.* 24, 175–219. doi: 10.1016/0010-0285(92)90007-O

Kanwisher, N. G. (1987). Repetition blindness: type recognition without token individuation. *Cognition* 27, 117–143. doi: 10.1016/0010-0277(87)90016-3

Keppel, G., and Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *J. Verbal Learning Verbal Behav.* 1, 153–161. doi: 10.1016/S0022-5371(62)80023-1

Kim, M.-S., and Cave, K. R. (1995). Spatial attention in visual search for features and feature conjunctions. *Psychol. Sci.* 6, 376–380. doi: 10.1111/j.1467-9280.1995.tb00529.x

Kleiner, M., Brainard, D., and Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception* 36, 1–16.

Klotz, W., and Neumann, O. (1999). Motor activation without conscious discrimination in metacontrast masking. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 976–992. doi: 10.1037/0096-1523.25.4.976

Krock, R. M., and Moore, T. (2015). The influence of gaze control on visual perception: eye movements and visual stability. *Cold Spring Harb. Symp. Quant. Biol.* 79, 123–130. doi: 10.1101/sqb.2014.79.024836

Kyllingsbæk, S. (2014). *Modeling Visual Attention: Encoding, Attention, and Short-Term Memory*. Doctoral dissertation, University of Copenhagen, Copenhagen.

LaRocque, J. J., Lewis-Peacock, J. A., and Postle, B. R. (2014). Multiple neural states of representation in short-term memory? It's a matter of attention. *Front. Hum. Neurosci.* 8:5. doi: 10.3389/fnhum.2014.00005

Loftus, G. R., and Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychon. Bull. Rev.* 1, 476–490. doi: 10.3758/BF03210951

Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Curr. Dir. Psychol. Sci.* 20, 240–245. doi: 10.1177/0963721411415340

Luck, S. J. (2008). "Visual short-term memory," in *Visual Memory*, eds S. J. Luck and A. Hollingworth (Oxford: Oxford University Press), 43–83.

Luck, S. J., and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature* 390, 279–281. doi: 10.1038/36846

Luck, S. J., and Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn. Sci.* 17, 391–400. doi: 10.1016/j.tics.2013.06.006

Ma, W. J., Husain, M., and Bays, P. M. (2014). Changing concepts of working memory. *Nat. Neurosci.* 17, 347–356. doi: 10.1038/nn.3655

Macmillan, N. A., and Creelman, C. D. (2005). *Detection Theory: A User's Guide*. New York, NY: Psychology Press.

Martens, S., and Wyble, B. (2010). The attentional blink: past, present, and future of a blind spot in perceptual awareness. *Neurosci. Biobehav. Rev.* 34, 947–957. doi: 10.1016/j.neubiorev.2009.12.005

Nosofsky, R. M., Little, D. R., Donkin, C., and Fific, M. (2011). Short-term memory scanning viewed as exemplar-based categorization. *Psychol. Rev.* 118, 280–315. doi: 10.1037/a0022494

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 411–421. doi: 10.1037//0278-7393.28.3.411

Oberauer, K., and Eichenberger, S. (2013). Visual working memory declines when more features must be remembered for each object. *Mem. Cogn.* 41, 1212–1227. doi: 10.3758/s13421-013-0333-6

Olivers, C. N. L., Peters, J., Houtkamp, R., and Roelfsema, P. R. (2011). Different states in visual working memory: when it guides attention and when it does not. *Trends Cogn. Sci.* 15, 327–334. doi: 10.1016/j.tics.2011.05.004

Palmeri, T. J., and Tarr, M. J. (2008). "Visual object perception and long-term memory," in *Visual Memory*, eds S. J. Luck and A. Hollingworth (Oxford: Oxford University Press), 163–207.

Pearson, B., Raskevicius, J., Bays, P. M., Pertzov, Y., and Husain, M. (2014). Working memory retrieval as a decision process. *J. Vis.* 14, 2. doi: 10.1167/14.2.2

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897X00366

Petersen, A., Kyllingsbæk, S., and Bundesen, C. (2012). Measuring and modeling attentional dwell time. *Psychon. Bull. Rev.* 19, 1029–1046. doi: 10.3758/s13423-012-0286-y

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Percept. Psychophys.* 16, 283–290. doi: 10.3758/BF03203943

Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25. doi: 10.1080/00335558008248231

Postle, B. R. (2015). The cognitive neuroscience of visual short-term memory. *Curr. Opin. Behav. Sci.* 1, 40–46. doi: 10.1016/j.cobeha.2014.08.004

Poth, C. H., Herwig, A., and Schneider, W. X. (2015). Breaking object correspondence across saccadic eye movements deteriorates object recognition. *Front. Syst. Neurosci.* 9:176. doi: 10.3389/fnsys.2015.00176

Poth, C. H., Petersen, A., Bundesen, C., and Schneider, W. X. (2014). Effects of monitoring for visual events on distinct components of attention. *Front. Psychol.* 5:930. doi: 10.3389/fpsyg.2014.00930

Poth, C. H., and Schneider, W. X. (2016). Breaking object correspondence across saccades impairs object recognition: the role of color and luminance. *J. Vis.* 16, 1–12. doi: 10.1167/16.11.1

R Development Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna: R foundation for statistical computing.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev.* 85, 59–108. doi: 10.1037/0033-295X.85.2.59

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., and Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* 16, 225–237. doi: 10.3758/PBR.16.2.225

Schneider, W. X. (1995). VAM: a neuro-cognitive model for visual attention control of segmentation, object recognition, and space-based motor action. *Vis. Cogn.* 2, 331–376. doi: 10.1080/13506289508401737

Schneider, W. X. (2013). Selective visual processing across competition episodes: a theory of task-driven visual attention and working memory. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368:20130060. doi: 10.1098/rstb.2013.0060

Shibuya, H., and Bundesen, C. (1988). Visual selection from multielement displays: measuring and modeling effects of exposure duration. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 591–600. doi: 10.1037/0096-1523.14.4.591

Soto, D., Mäntylä, T., and Silvanto, J. (2011). Working memory without consciousness. *Curr. Biol.* 21, R912–R913. doi: 10.1016/j.cub.2011.09.049

Sperling, G. (1960). The information available in brief visual presentations. *Psychol. Monogr.* 74, 1–29. doi: 10.1037/h0093759

Sternberg, S. (1966). High-speed scanning in human memory. *Science* 153, 652–654. doi: 10.1126/science.153.3736.652

Sternberg, S. (1975). Memory scanning: new Findings and current controversies. *Q. J. Exp. Psychol.* 27, 1–32. doi: 10.1080/14640747508400459

Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn. Sci.* 19, 394–405. doi: 10.1016/j.tics.2015.05.004

Strasburger, H., Rentschler, I., and Jüttner, M. (2011). Peripheral vision and pattern recognition: a review. *J. Vis.* 11, 13. doi: 10.1167/11.5.13

Treisman, A., and Gelade, G. (1980). A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136. doi: 10.1016/0010-0285(80)90005-5

Wheeler, M. E., and Treisman, A. M. (2002). Binding in short-term visual memory. *J. Exp. Psychol. Gen.* 131, 48–64. doi: 10.1037/0096-3445.131.1.48

Wickelgren, W. A. (1970). Time, interference, and rate of presentation in short-term recognition memory for items. *J. Math. Psychol.* 7, 219–235. doi: 10.1016/0022-2496(70)90045-3

Wilken, P., and Ma, W. J. (2004). A detection theory account of change detection. *J. Vis.* 4, 1120–1135. doi: 10.1167/4.12.11

Zhou, F., Kahana, M. J., and Sekuler, R. (2004). Short-term episodic memory for visual textures: a roving probe gathers some memory. *Psychol. Sci.* 15, 112–118. doi: 10.1111/j.0963-7214.2004.01502007.x

Study 5

Prioritization in visual working memory enhances memory-retention and speeds up processing in a comparison task

(submitted to Attention, Perception, and Psychophysics)

Prioritization in visual working memory enhances memory-retention and speeds up

processing in a comparison task

Christian H. Poth & Werner X. Schneider

Bielefeld University

Author note

Christian H. Poth, Neuro-cognitive Psychology, Department of Psychology, and

Cluster of Excellence Cognitive Interaction Technology, Bielefeld University. P.O. box

100131, 33501 Bielefeld, Germany.

Werner X. Schneider, Neuro-cognitive Psychology, Department of Psychology, and

Cluster of Excellence Cognitive Interaction Technology, Bielefeld University. P.O. box

100131, 33501 Bielefeld, Germany.

Correspondence concerning this article should be addressed to Christian H. Poth,

Neuro-cognitive Psychology, Department of Psychology, and Cluster of Excellence Cognitive

Interaction Technology, Bielefeld University. P.O. box 100131, 33501 Bielefeld, Germany.

E-mail: c.poth@uni-bielefeld.de

Author contributions: CHP conceived the idea of this project, developed the

experimental paradigm, programmed the experiments, analyzed the data, interpreted the

results, and wrote the manuscript. WXS discussed the idea of this project, the development of

the experimental paradigm, interpreted the results, and revised the manuscript.

Words: 4595, figures: 3, tables: 1

Abstract

Visual working memory retains visual information for controlling behavior. We studied how information in visual working memory is prioritized for being used. In two experiments, participants memorized the stimuli of a memory display for a brief interval, followed by a retro-cue. The retro-cue was either valid, indicating which stimulus from the memory display was relevant (i.e. had priority) in the upcoming comparison with a probe, or was neutral (uninformative). Next, the probe was presented, terminated by a mask, and participants reported whether it matched a stimulus from the memory display. The presentation duration of the probe was varied. Assessing performance as a function of presentation duration allowed to disentangle two components of working memory: memory-retention and the speed of processing the probe for the memory-based comparison. Compared with neutral retro-cues, valid retro-cues improved retention and at the same time accelerated processing of the probe. These findings show for the first time that prioritization in working memory impacts on distinct mechanisms: Retrospectively, it supports memory-retention, and prospectively, it enhances perceptual processing in upcoming comparison tasks.

Prioritization in visual working memory enhances memory-retention and speeds up

processing in a comparison task

Visual working memory (VWM) is a cornerstone of human visual cognition. It

temporarily retains visual information and makes it accessible for cognitive operation, report,

and action control (Eriksson, Vogel, Lansner, Bergstrom, & Nyberg, 2015; Oberauer, 2009;

Poth & Schneider, 2016a; Schneider, 2013). VWM has only limited capacity (Luck & Vogel,

1997; Shibuya & Bundesen, 1988; Sperling, 1960). Efficient use of this capacity dictates

selectivity: Currently relevant information must enter VWM with priority over less relevant

information. This prioritization is performed by mechanisms of visual attention (Bundesen,

1990; Bundesen, Habekost, & Kyllingsbæk, 2005; Duncan & Humphreys, 1989; Schneider,

1995). The bulk of attention research focused on prioritization up to the time of encoding into

VWM (Bundesen, Vangkilde, & Petersen, 2015; Duncan, 2006; Poth & Schneider, 2013).

However, flexible visual cognition requires that changes of priority can be accommodated

also when they happen after information has entered VWM.

Indeed, more recent research demonstrated that prioritization continues after encoding

into VWM. This research made use of the retro-cuing paradigm (Griffin & Nobre, 2003;

Landman, Spekreijse, & Lamme, 2003). Participants memorized a set of visual stimuli, the

memory display, over a retention interval which was followed by a probe stimulus. The task

was to report whether the probe matched an item from the memory display. A so-called *retro-*

cue (i.e. a "retrodictive" cue) was shown after the memory display but before the probe

appeared. In the experiments of current interest, retro-cues could be valid or neutral (Astle,

Summerfield, Griffin, & Nobre, 2012; Kuo, Stokes, & Nobre, 2012). A valid retro-cue

predicted which of the items from the memory display was going to be relevant for the

upcoming comparison with the probe. A neutral retro-cue did not contain any predictive

information regarding this comparison. The central result is that valid retro-cues improved

comparison performance relative to neutral retro-cues. Over a decade of research accumulated

evidence for beneficial effects of valid retro-cues in different versions of the basic paradigm (Astle et al., 2012; Griffin & Nobre, 2003; Landman et al., 2003; Makovski & Jiang, 2007; Makovski, Sussman, & Jiang, 2008; Souza, Rerko, & Oberauer, 2014). Thus, it seems safe to conclude that valid retro-cues prioritize an item from a preceding memory display while the memory display is retained in VWM.

Still controversial, however, is the question which mechanisms underlie the prioritization within VWM (Souza & Oberauer, 2016). Current accounts assume that valid retro-cues improve comparison performance by manipulating the representations of the memory display in VWM. Specifically, some authors propose that they strengthen the VWM representation of the cued item, increasing the utility of this item for the comparison (Kuo, Yeh, Chen, & D'Esposito, 2011; Lepsien, Thornton, & Nobre, 2011; Nobre, Griffin, & Rao, 2008). Others propose that they free VWM capacity and reduce interference within VWM by having uncued items removed from VWM (Souza et al., 2014; Williams, Hong, Kang, Carlisle, & Woodman, 2013). Again others suggest that valid retro-cues protect the cued item against decay (Matsukura, Luck, & Vecera, 2007; Murray, Nobre, Clark, Cravo, & Stokes, 2013; Pertzov, Bays, Joseph, & Husain, 2013) or new interfering information (such as from the probe; Makovski et al., 2008; Makovski & Jiang, 2007). Finally, some suggest that valid retro-cues grant the cued items priority in the process of being compared to the probe (Astle et al., 2012; Makovski et al., 2008; Nobre et al., 2008). Fundamental to all these accounts is that retro-cues are, as the term implies, retroactive. That is, all accounts assume that valid retro-cues engage mechanisms that, in one way or the other, prioritize information from the past which is now retained in VWM.

Here, we ask whether retro-cues facilitate memory-retention in VWM, or whether they enhance future perceptual processing in service of the comparison task, or both. To this end, we introduce a novel paradigm which allows to disentangle such retrospective and prospective effects of retro-cues (Figure 1). Participants briefly viewed a memory display of

two colored squares and memorized them over a retention interval. This interval outlasted

iconic memory traces and thus called for retention in VWM (for a review, see Irwin &

Thomas, 2008). Afterwards, a probe stimulus was presented that either matched or did not

match an item from the memory display, with each alternative occurring in half the trials.

Participants' task was to indicate whether or not the probe matched an item from the memory

display. A retro-cue appeared after the retention interval but before the probe. The retro-cue

was either valid or neutral. A valid retro-cue pointed at the location of one of the items from

the preceding memory display, the one that was going to be relevant for the upcoming

comparison with the probe. Specifically, if the probe matched an item from the memory

display, it was always the one indicated by the retro-cue. A neutral retro-cue did not contain

any specific location information. We varied the presentation duration of the probe and

terminated it with a pattern mask. This enabled us to assess performance as a function of the

presentation duration. To disentangle the retrospective and the prospective effects of retro-

cues, we fit these data with an exponential model (cf. Bundesen, 1990; Wickelgren, 1977) and

compared the estimated parameters between valid and neutral retro-cues. The exponential

model comprised three parameters. First, the level of asymptotic performance which is

reached when the probe is shown long enough (see the upper asymptotes of the smooth curves

in Figure 2). The perceptual processing of the probe should improve with increasing

presentation duration (e.g., Petersen & Andersen, 2012; Shibuya & Bundesen, 1988). At the

asymptote, however, performance stops to increase with increasing presentation duration of

the probe. Therefore, when the asymptote were reached, perceptual processing (encoding) of

the probe should be over and variations of the asymptote should be due to post-perceptual

factors. In the present case, variations in the asymptote should reflect the performance level

for retaining the items of the memory display in VWM. Second, the rate at which

performance increases with increasing presentation duration of the probe toward asymptotic

performance (see how steeply the smooth curves increase with increasing presentation

duration in Figure 2). This is a well-established measure of processing speed (Bundesen, 1990; Bundesen et al., 2015; Shibuya & Bundesen, 1988; Wickelgren, 1977). Applied to the present case, it should represent the speed with which the probe is perceptually processed in order to accomplish the comparison with the items of the memory display. Third, a temporal threshold, reflecting the presentation duration of the probe that must be exceeded to increase performance above chance level (cf. Bundesen, 1990; Wickelgren, 1977).

If retro-cues facilitate memory-retention in VWM, then valid retro-cues should result in a higher asymptotic performance than neutral ones. In contrast, if retro-cues enhance processing for the upcoming comparison task, then valid retro-cues should lead to a higher processing speed of the probe than neutral ones. We tested these hypotheses in Experiment A and, in addition, performed a close replication with Experiment B.

Method

Participants

Eleven participants were paid to perform Experiment A. One additional participant was excluded because of problems with the eye-tracker calibration (see below) and corrupted eye-tracking data files. The participants were between 21 and 34 years old ($MD = 24$ years), eight were female, three were male, and all reported normal or corrected-to-normal (contact lenses) visual acuity and normal color vision. Thirteen participants received course credit for performing Experiment B. They were between 18 and 27 years old ($MD = 20$ years), eleven were female, two male, and all reported normal or corrected-to-normal visual acuity and normal color vision. All participants gave written informed consent before participation, the experiments were approved by Bielefeld University's ethics committee, and complied with the ethical guidelines of the German Psychological Association (DGPs).

Apparatus and stimuli

Both experiments took place in a dimly lit room. A chin-and-forehead rest (Experiment A), and a chin rest (Experiment B) ensured that participants viewed the CRT-

monitors from a distance of 71 cm. In Experiment A, the CRT-monitor (G90fB, ViewSonic, Brea, CA, USA) ran at a refresh rate of 100 Hz and a resolution of 1024×768 pixels (corresponding to physical dimensions of 36×27 cm). For control purposes, a video-based desktop-mounted eye tracker sampled participants' right eyes at 1000 Hz (Eyelink 1000, SR Research, Mississauga, Ontario, Canada; 9-point grid calibration). In Experiment B, the CRT-monitor (Trinitron MultiScan G420, Sony, Park Ridge, NJ, USA) ran at a refresh rate of 85 Hz and a resolution of 1024×768 pixels (corresponding to physical dimensions of 36×27 cm). Responses were collected using standard computer keyboards (QWERTZ-layout). The experiments were controlled by the Psychophysics toolbox (3.0.12; Kleiner et al., 2007; Pelli, 1997, and in Experiment A, the Eyelink toolbox, 3.0.12; Cornelissen, Peters, & Palmer, 2002) extensions for Matlab (R2014b in Experiment A, R2013b in Experiment B, The MathWorks, Natick, MA, USA).

Stimulus-luminance was measured using an LS-110 luminance meter (MINOLTA, Osaka, Japan). Stimulus-luminance is reported for Experiment A and B side by side (i.e. luminance A/luminance B). An empty light-gray square was used as fixation stimulus (0.67×0.67° of visual angle, with a linewidth of 4 pixels; 59/45 cd/m^2), henceforth called fixation square. Valid retro-cues consisted in a brightening of two lines of this square, neutral ones in a brightening of the whole square (102/114 cd/m^2). The stimuli of the memory display were squares (0.67×0.67°) of the following eight colors: red (RGB: 255, 0, 0; 34/23 cd/m^2), magenta (RGB: 255, 0, 255; 43/34 cd/m^2), yellow (RGB: 255, 255, 0; 112/101 cd/m^2), orange (RGB: 255, 145, 0; 59/44 cd/m^2), blue (RGB: 0, 0, 255; 17/12 cd/m^2), cyan (RGB: 0, 255, 255; 103/92 cd/m^2), green (RGB: 0, 255, 0; 100/80 cd/m^2), and black (RGB: 0, 0, 0; 1/< 1 cd/m^2). For each individual participant, 99 pattern masks were algorithmically created in the beginning of the experiment. Masks consisted of a square composed of a 4×4 matrix of smaller squares (0.30×0.30° each) whose colors were randomly chosen from the set of colors with the constraint that each color occurred twice in each mask (see, Figure 1 for an example

mask). The gray background had a luminance of 34/22 cd/m^2. The white question mark of the

response screen was written in Arial (0.50°×1.00°; 102/114 cd/m^2).

Procedure and design

Figure 1 illustrates the paradigm of Experiment A (the paradigm of Experiment B was

similar in most respects, see below). Participants started each trial by pressing the space-bar.

In the beginning of a trial, the fixation square was shown at screen center for 400 ms. In

Experiment A, the eye-tracker monitored if eyes were open (i.e. pupils visible) until the end

of this fixation period and if they were not, the period was prolonged until the next screen

refresh after the eyes were open again. In Experiment B, eye behavior was not recorded. The

fixation square stayed on for the most of a trial. After the fixation period, the memory display

containing two differently colored squares was shown for 100 ms in Experiment A, and for 94

ms in Experiment B. The colors of the squares were randomly chosen from the set of used

colors. The squares appeared at two out of four possible positions (2° from screen center

horizontally to the left or right × vertically to the left or right, see Figure 1) and this was

randomized across trials with each pairing of positions occurring equally often. The memory

display was followed by an interstimulus interval (ISI) of 1000 ms, after which the retro-cue

was presented for 100 ms in Experiment A and 94 ms in Experiment B. Valid retro-cues

consisted in a selective brightening of two sides of the fixation square, forming an arrow

pointing at one of the two location of the squares of the previous memory display. Across

trials, each location was cued equally often. Neutral retro-cues consisted in a brightening of

the whole fixation square. The retro-cue was followed by another ISI of 1000 ms. The probe

then replaced the fixation square at screen center. In Experiment A, the probe was shown for

eight different durations (10, 20, 30, 40, 60, 90, 120, and 180 ms). In Experiment B, it was

shown for six different durations (12, 24, 35, 59, 106, and 141 ms). The presentation of the

probe was terminated by a central pattern mask lasting for 300 ms in Experiment A and 306

ms in Experiment B. Afterwards, a central question mark was presented until participants

responded. Participants were instructed to respond with the F9-key if the probe matched an item from the memory display and the F1-key if it did not match any of them. There was no time limit for the response. On half of the trials, the color of the probe matched the color of one of the items from the memory display. On the other half, it was had one of the colors that did not appear on this trial. Participants were informed that if a valid retro-cue was shown and the probe matched the color of an item from the memory display, then this would be the color of the item indicated by the retro-cue.

Participants performed Experiment A in a single session of 768 trials, 48 trials per retro-cue condition (valid vs. neutral) and per presentation duration of the probe. Participants performed Experiment B in two sessions (on separate days) of 576 trials each, yielding 1152 trials in total, 96 trials per retro-cue condition and per presentation duration of the probe. Trials were administered in randomized order in Experiment A and in both sessions of Experiment B. Participants performed 30 training trials (randomly chosen with replacement from all trial types) before each participation.
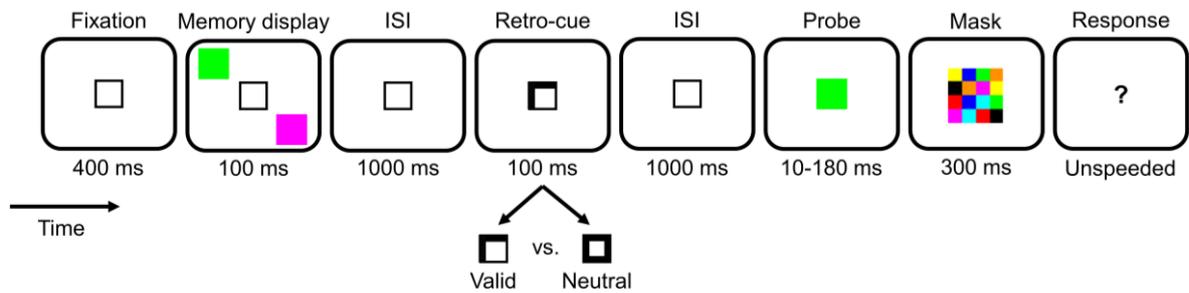
Figure 1. Paradigm of Experiment A. The paradigm of Experiment B was similar but differed in the display durations (and other aspects, see the Methods). Participants fixated a fixation square, after which a memory display with two colored squares appeared. After an interstimulus interval (ISI), a retro-cue was shown. If the retro-cue was valid, it indicated the location of the previous item from the memory display that was going to be relevant for the current trial. If it was neutral, it did not contain any predictive information. After another ISI, a probe was presented whose presentation duration was parametrically varied across trials. The probe was terminated by a pattern mask. In the end of a trial, a question mark prompted participants to indicate whether or not the probe matched an item from the memory display.

Results

To control for response biases, performance in indicating whether the probe matched an item from the memory display was assessed as d' (the z-transformed rate of "yes"-responses to probes matching an item from the memory display minus the z-transformed rate of "yes"-responses to probes not matching an item from the memory display; 0.5 was added to the data cells on which rates were computed to prevent infinite values of d', see Macmillan & Creelman, 2005). Performance was assessed as a function of presentation duration of the probe. For each participant and each retro-cue condition, these data were fit with an exponential model of the type

$$d' = \omega\left(1 - \exp\left(-v * (t - t_0)\right)\right)$$

where ω is the upper asymptote of performance (in d'), v is the rate parameter of the exponential distribution which measures processing speed of the probe within the comparison with the items of the memory display (in items/s; cf. Bundesen, 1990; Bundesen et al., 2015; Wickelgren, 1977). The third parameter, t_0, is a temporal threshold consisting in the presentation duration of the probe that has to be exceeded to increase performance above chance level (in ms; cf. Bundesen, 1990). Fitting was accomplished using the nonlinear least-squares (nls) method implemented in R (3.0.3; R Core Team, 2016). Figure 2 depicts the performance in each retro-cue condition and corresponding model fits, for one participant in each experiment. Table 1 provides descriptive statistics of estimated parameters for the two retro-cue conditions of both experiments. Goodness-of-fit was quantified as Pearson's correlation r between the predicted values based on the fitted model and participants' observed values (see, Table 1).
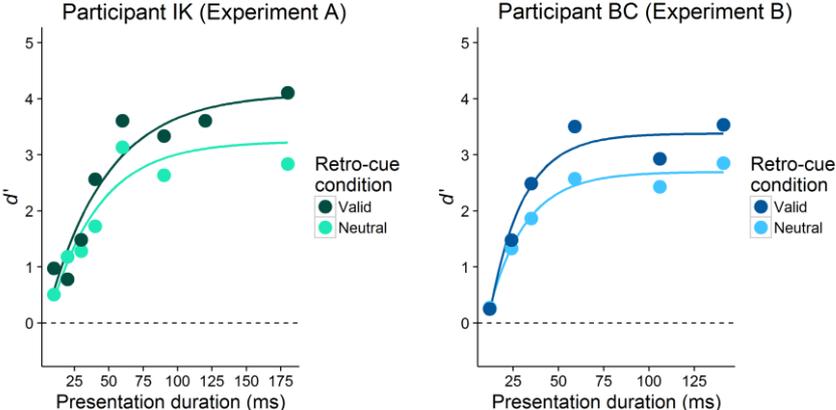
Figure 2. Performance of a participant of Experiment A (left) and one of Experiment B (right). Data points

represent performance in indicating whether or not probes matched an item from the memory display at each of

the presentation durations of the probe. The two retro-cue conditions are depicted separately. Smooth curves

represent least-squares fits of the exponential model to the data of the two retro-cue conditions.

The parameter estimates were compared between the retro-cue conditions using paired-samples *t*-tests (two-sided, with a significance criterion of p < .05, and d_z as effect size Cohen, 1988, and for which the assumption of normally distributed differences was assessed beforehand using Shapiro-Wilk tests). *t*-tests were supplemented with corresponding Bayes Factors (*BF*$_{10}$, Rouder, Speckman, Sun, Morey, & Iverson, 2009, whereby values greater one favor the alternative and values smaller one favor the null hypothesis). Figure 3 illustrates the parameter comparisons between the retro-cue conditions of the two experiments.

In Experiment A, the retention performance ω was significantly higher when retro-cues were valid than when they were neutral, $t(10) = 3.212$, $p = .009$, $d_z = 0.97$, $BF_{10} = 6.42$. Experiment B replicated this result, $t(12) = 2.808$, $p = .016$, $d_z = 0.78$, $BF_{10} = 3.99$. These findings demonstrate a retrospective effect of retro-cues. They indicate that valid retro-cues facilitated the retention of the relevant item from the memory display as compared with neutral retro-cues.

The processing speed *v* was significantly higher when retro-cues were valid compared with neutral, and this was likewise the case in Experiment A, $t(10) = 2.400$, $p = .037$, $d_z = 0.72$, $BF_{10} = 2.14$, and Experiment B, $t(12) = 2.589$, $p = .024$, $d_z = 0.72$, $BF_{10} = 2.89$. These findings reveal a prospective effect of retro-cues on future visual processing. Specifically, compared with neutral retro-cues, valid retro-cues increased the speed with which an upcoming probe was processed, in order to assess whether it had been contained in the preceding memory display.

There were no differences between the valid and the neutral retro-cue condition regarding the temporal threshold, t_0, neither in Experiment A, $t(10) = -0.629$, $p = .543$, $d_z = -0.19$, $BF_{10} = 0.35$, nor in Experiment B, $t(12) = 0.787$, $p = .447$, $d_z = 0.22$, $BF_{10} = 0.36$.
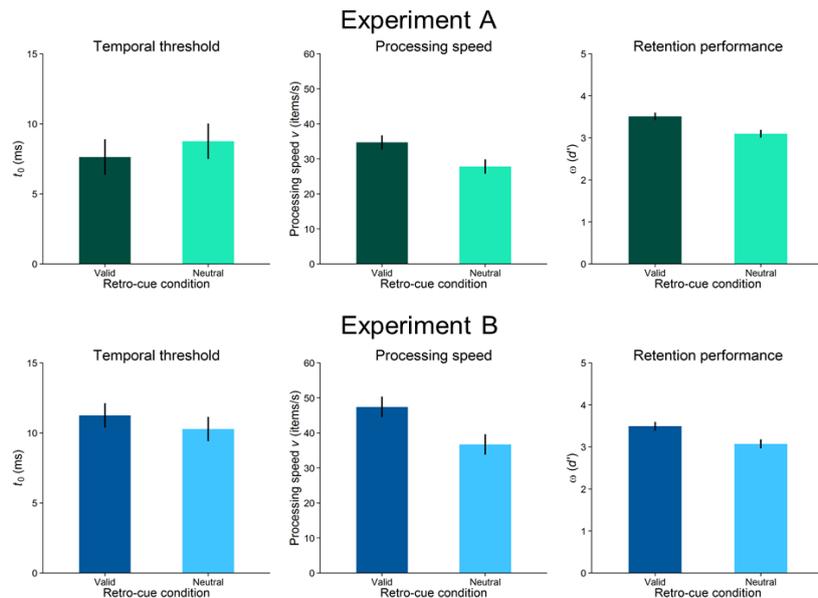
Figure 3. Estimated model parameters in the two retro-cue conditions of Experiment A (upper panel) and Experiment B (lower panel). Depicted are the means of the retention performance ω (in *d'*), of the processing speed *v* (in items/s), and of the temporal threshold t_0 (in ms), across participants. Error-bars indicate ± one standard error for within-subjects designs (Loftus & Masson, 1994).

Table 1. Descriptive statistics of Experiment A and B.

	Experiment A		Experiment B	
	Valid retro-cue	Neutral retro-cue	Valid retro-cue	Neutral retro-cue
ω	3.51 (0.50)	3.10 (0.64)	3.49 (0.81)	3.07 (0.89)
v	34.70 (11.08)	27.80 (10.53)	47.42 (20.22)	36.72 (14.72)
t_0	8 (3)	9 (5)	11 (3)	10 (3)
r	.945 (.030)	.917 (.036)	.962 (.027)	.961 (.029)

Note. Descriptive statistics of estimated parameters for the two retro-cue conditions of Experiment A and B. Provided are means and standard deviations (in parentheses) across participants for the asymptotic performance level ω (in *d'*), the processing speed *v* (in items/s), the temporal threshold t_0 (in ms), and for Pearson's correlation between the values predicted by the fitted model and the observed values.

Discussion

The present experiments reveal that retro-cues impact on two distinct components of VWM performance. First, valid retro-cues improve memory-retention in VWM. This was evident from the higher retention performance when retro-cues were valid compared with neutral. We assessed retention performance as the asymptote of performance as a function of the presentation duration of the probe. Asymptotic performance reflects a component of VWM performance that is independent from perceptual processing of the probe, because performance does not improve further when the probe is presented longer. Second, valid retro-cues enhanced the speed with which probe stimuli were perceptually processed in order to be compared to the items retained in VWM. We assessed the speed of processing the probe for this memory-based comparison as the rate of performance increase with increasing presentation duration of the probe.

Valid retro-cues improved memory-retention. This indicates that prioritization modulates representations in VWM independently of the time available to process the probe for the upcoming comparison task. This is in line with several not mutually exclusive accounts assuming that retro-cues impact on VWM-based performance by directly modulating VWM representations. Valid retro-cues may strengthen the representations of cued items in VWM (Kuo et al., 2011; Lepsien et al., 2011; Nobre et al., 2008). They may remove uncued items from VWM, thereby freeing capacity and reducing interference within VWM (Souza et al., 2014; Williams et al., 2013). They may protect the representations of cued items in VWM against decay (Matsukura et al., 2007; Pertzov et al., 2013) or interfering new information (Makovski et al., 2008; Makovski & Jiang, 2007).

Crucially, valid retro-cues also accelerated the perceptual processing of probe stimuli for performing the upcoming comparison task. This indicates that prioritization in VWM has effects beyond pure memory-retention. In this way, this present finding calls for an extension

of current accounts of the beneficial effects of valid retro-cues on VWM-based performance. The finding can be interpreted in two ways.

The first explanation is that the presentation duration of the probe stimulus determined the quality of its representation in VWM. The rate of performance increase with increasing probe presentation duration was higher when retro-cues were valid than neutral. Thus, a valid retro-cues may have compensated for the low representational quality of the probe at short presentation durations, for example by enhancing the VWM representations of the cued items, which would be in line with previous accounts (e.g., Kuo et al., 2011; Lepsien et al., 2011; Nobre et al., 2008). The present findings would then show that valid retro-cues improve a component of VWM that can be traded-off for the representational quality of the probe stimulus. However, with the effects on memory-retention, the findings would also indicate that valid retro-cues improve a VWM component beyond this, one that does not interact with the representational quality of the probe.

The second explanation assumes that the acceleration of the processing of the probe resulted from a prospective monitoring process. A valid retro-cue indicates which of the items in VWM will be relevant for the comparison to an upcoming probe. Consequently, the environment may be monitored for the features of the cued item already before the probe appears. It has previously been shown that monitoring processes rely on visual attention (Poth, Petersen, Bundesen, & Schneider, 2014). In the present case, the features of the cued item may be monitored for by engaging the pigeonholing ("attention to features") mechanism from Bundesen' theory of visual attention (TVA; Bundesen, 1990; Bundesen et al., 2015). Pigeonholing influences the speed with which visual features of objects are processed for being encoded into VWM. This should happen by up- or down-regulating a perceptual bias for categorizing any given object as having a certain feature. The perceptual bias is internal, meaning that it is independent of what objects are actually viewed. Specifically, the perceptual bias for a certain feature may be implemented by increasing or decreasing the

firing rates of visual neurons preferentially coding for this feature (Bundesen et al., 2005). Increasing the perceptual bias for the features of the cued item would increase the speed of encoding the probe into VWM, if the probe has these features. This would imply that the present processing acceleration stemmed from the trials on which the probe matched the cued item. Indeed, effects of retro-cues have been found more pronounced for such matches (Lepsien, Griffin, Devlin, & Nobre, 2005; Nobre et al., 2008). However, a processing acceleration on trials on which the probe did not match an item from the memory display could still be explained in terms of pigeonholing by assuming an additional decision deadline (cf. Bundesen's 1990 explanation of target-absent response times in visual search). In this scenario, valid retro-cues would lower the deadline for processing the probe in order for deciding that it did not match an item from the memory display.

On the neuronal level, active or passive working memory processes may underlie the effects of valid retro-cues. Both types of processes rely on the sensory recruitment hypothesis, stating that the retention of information in VWM is enacted by the same visual brain areas that encode this information at first (Ester, Anderson, Serences, & Awh, 2013; Miller, Erickson, & Desimone, 1996; Serences, Ester, Vogel, & Awh, 2009; Supèr, Spekreijse, & Lamme, 2001). The hypothesis of active working memory processes relies on the assumption that retention in VWM is performed by maintaining the spiking activity of the neurons coding for the retained items and their features (Chelazzi, Miller, Duncan, & Desimone, 1993). The retrospective effects of valid retro-cues on memory-retention in VWM could consist in an increase of activity in neurons coding for the features of the cued item, or a decrease in those coding for features of other items, or both (Lepsien et al., 2011; Trapp & Lepsien, 2012). An increase of spiking activity of these neurons could at the same time provoke the prospective effects of valid retro-cues. The increased firing would support the future encoding of stimuli by these neurons, explaining why processing of the probe was accelerated in the present experiments.

The hypothesis of passive working memory processes is grounded on recent evidence questioning maintained spiking activity as the sole neuronal implementation of retention in VWM (for an overview, see Stokes, 2015). Findings from single-cell neurophysiology (Stokes et al., 2013) and brain imaging and stimulation (e.g., Rose et al., 2016; for a review, see Larocque, Lewis-Peacock, & Postle, 2014) suggest a passive (V)WM (Schneider, 2013), which presumably relies on increased synaptic connectivity rather than maintained spiking activity (Mongillo, Barak, & Tsodyks, 2008; though synaptic connectivity may stem from an initial increase of neuronal firing, e.g., Stokes, 2015). Valid retro-cues could also exert their effects here, by modulating the synaptic connectivity in visual brain areas at short time scales. Valid retro-cues could increase the synaptic connectivity of neurons coding for the features of the cued item. Enhanced synaptic connectivity could make VWM representations more robust, offering another explanation why valid retro-cues improved retention performance. Interestingly, if valid retro-cues increased the short-term synaptic connectivity of these neurons, this might also increase their efficiency of encoding new stimuli (Stokes et al., 2013; Sugase-Miyamoto, Liu, Wiener, Optican, & Richmond, 2008). Hence, this provides another explanation of why valid retro-cues accelerated processing of probes. It is important to note, however, that both, active and passive working memory processes are feature-based. Therefore, both processes would only be able to operate if the features of probes matched the features of the cued items in VWM. This would call for an additional process implementing a decision deadline, in the same way as the prospective monitoring process that relied on Bundesen's (1990) pigeonholing mechanism.

To conclude, the present study shows that priority within VWM not only affects the retention of past information but also future processing in a comparison task. Visual attention seems not only to set processing priorities for encoding into VWM (Bundesen, 1990) and for selection within VWM (Griffin & Nobre, 2003; Landman et al., 2003), but at the same time also for processing upcoming information. In this vein, visual attention may provide a bridge

between episodes of visual processing that are composed of encoding into and retention in

VWM (as proposed by Schneider, 2013; see also Poth, Herwig, & Schneider, 2015; Poth &

Schneider, 2016b).

References

Astle, D. E., Summerfield, J., Griffin, I., & Nobre, A. C. (2012). Orienting attention to locations in mental representations. *Attention, Perception & Psychophysics*, *74*(1), 146–162. doi:10.3758/s13414-011-0218-3

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523–547. doi:10.1037/0033-295X.97.4.523

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review*, *112*(2), 291–328. doi:10.1037/0033-295X.112.2.291

Bundesen, C., Vangkilde, S., & Petersen, A. (2015). Recent developments in a computational theory of visual attention (TVA). *Vision Research*, *116*, 210–218. doi:10.1016/j.visres.2014.11.005

Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, *363*(6427), 345–347. doi:10.1038/363345a0

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 613–617. doi:10.3758/BF03195489

Duncan, J. (2006). EPS Mid-Career Award 2004: Brain mechanisms of attention. *Quarterly Journal of Experimental Psychology*, *59*(1), 2–27. doi:10.1080/17470210500260674

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458. doi:10.1037/0033-295X.96.3.433

Eriksson, J., Vogel, E. K., Lansner, A., Bergstrom, F., & Nyberg, L. (2015). Neurocognitive architecture of working memory. *Neuron*, *88*(1), 33–46. doi:10.1016/j.neuron.2015.09.020

Ester, E. F., Anderson, D. E., Serences, J. T., & Awh, E. (2013). A neural measure of precision in visual working memory. *Journal of Cognitive Neuroscience*, *25*(5), 754–761. doi:10.1162/jocn_a_00357

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, *15*(8), 1176–1194. doi:10.1162/089892903322598139

Irwin, D. E., & Thomas, L. E. (2008). Visual sensory memory. In S. J. Luck & A. Hollingworth (Eds.), *Visual memory* (pp. 9–41). New York: Oxford University Press.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception*, *36*(14), 1.

Kuo, B.-C., Stokes, M. G., & Nobre, A. C. (2012). Attention modulates maintenance of representations in visual short-term memory. *Journal of Cognitive Neuroscience*, *24*(1), 51–60. doi:10.1162/jocn_a_00087

Kuo, B.-C., Yeh, Y.-Y., Chen, A. J.-W., & D'Esposito, M. (2011). Functional connectivity during top-down modulation of visual short-term memory representations. *Neuropsychologia*, *49*(6), 1589–1596. doi:10.1016/j.neuropsychologia.2010.12.043

Landman, R., Spekreijse, H., & Lamme, V. A. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research*, *43*(2), 149–164. doi:10.1016/S0042-6989(02)00402-9

Larocque, J. J., Lewis-Peacock, J. A., & Postle, B. R. (2014). Multiple neural states of representation in short-term memory? It's a matter of attention. *Frontiers in Human Neuroscience*, *8*, 5. doi:10.3389/fnhum.2014.00005

Lepsien, J., Griffin, I. C., Devlin, J. T., & Nobre, A. C. (2005). Directing spatial attention in mental representations: Interactions between attentional orienting and working-memory load. *NeuroImage*, *26*(3), 733–743. doi:10.1016/j.neuroimage.2005.02.026

Lepsien, J., Thornton, I., & Nobre, A. C. (2011). Modulation of working-memory maintenance by directed attention. *Neuropsychologia*, *49*(6), 1569–1577. doi:10.1016/j.neuropsychologia.2011.03.011

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*(4), 476–490. doi:10.3758/BF03210951

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. doi:10.1038/36846

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Makovski, T., & Jiang, Y. V. (2007). Distributing versus focusing attention in visual short-term memory. *Psychonomic Bulletin & Review*, *14*(6), 1072–1078. doi:10.3758/BF03193093

Makovski, T., Sussman, R., & Jiang, Y. V. (2008). Orienting attention in visual working memory reduces interference from memory probes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(2), 369–380. doi:10.1037/0278-7393.34.2.369

Matsukura, M., Luck, S. J., & Vecera, S. P. (2007). Attention effects during visual short-term memory maintenance: Protection or prioritization? *Perception & Psychophysics*, *69*(8), 1422–1434. doi:10.3758/BF03192957

Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, *16*(16), 5154–5167.

Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic theory of working memory. *Science*, *319*(5869), 1543–1546. doi:10.1126/science.1150769

Murray, A. M., Nobre, A. C., Clark, I. A., Cravo, A. M., & Stokes, M. G. (2013). Attention restores discrete items to visual short-term memory. *Psychological Science*, *24*(4), 550–556. doi:10.1177/0956797612457782

Nobre, A. C., Griffin, I. C., & Rao, A. (2008). Spatial attention can bias search in visual short-term memory. *Frontiers in Human Neuroscience*, *1*, 4. doi:10.3389/neuro.09.004.2007

Oberauer, K. (2009). Design for a working memory. *Psychology of Learning and Motivation*, *51*, 45–100. doi:10.1016/S0079-7421(09)51002-X

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. doi:10.1163/156856897X00366

Pertzov, Y., Bays, P. M., Joseph, S., & Husain, M. (2013). Rapid forgetting prevented by retrospective attention cues. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), 1224–1231. doi:10.1037/a0030947

Petersen, A., & Andersen, T. S. (2012). The effect of exposure duration on visual character identification in single, whole, and partial report. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(2), 498–514. doi:10.1037/a0026728

Poth, C. H., Herwig, A., & Schneider, W. X. (2015). Breaking object correspondence across saccadic eye movements deteriorates object recognition. *Frontiers in Systems Neuroscience*, *9*, 176. doi:10.3389/fnsys.2015.00176

Poth, C. H., Petersen, A., Bundesen, C., & Schneider, W. X. (2014). Effects of monitoring for visual events on distinct components of attention. *Frontiers in Psychology*, *5*, 930. doi:10.3389/fpsyg.2014.00930

Poth, C. H., & Schneider, W. X. (2013). Aufmerksamkeit [Attention]. In A. Stephan & S. Walter (Eds.), *Handbuch Kognitionswissenschaft* (pp. 221–230). Weimar: J. B. Metzler.

Poth, C. H., & Schneider, W. X. (2016a). Episodic short-term recognition requires encoding into visual working memory: Evidence from probe recognition after letter report. *Frontiers in Psychology*, *7*, 1440. doi:10.3389/fpsyg.2016.01440

Poth, C. H., & Schneider, W. X. (2016b). Breaking object correspondence across saccades impairs object recognition: The role of color and luminance. *Journal of Vision*, *16*(11), 1. doi:10.1167/16.11.1

R Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, *354*(6316), 1136–1139. doi:10.1126/science.aah7011

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. doi:10.3758/PBR.16.2.225

Schneider, W. X. (1995). VAM: A neuro-cognitive model for visual attention control of segmentation, object recognition, and space-based motor action. *Visual Cognition*, *2*(2-3), 331–376. doi:10.1080/13506289508401737

Schneider, W. X. (2013). Selective visual processing across competition episodes: a theory of task-driven visual attention and working memory. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(1628), 20130060. doi:10.1098/rstb.2013.0060

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, *20*(2), 207–214. doi:10.1111/j.1467-9280.2009.02276.x

Shibuya, H., & Bundesen, C. (1988). Visual selection from multielement displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(4), 591–600. doi:10.1037/0096-1523.14.4.591

Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception & Psychophysics*, *78*(7), 1839–1860. doi:10.3758/s13414-016-1108-5

Souza, A. S., Rerko, L., & Oberauer, K. (2014). Unloading and reloading working memory: attending to one item frees capacity. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1237–1256. doi:10.1037/a0036331

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1–29. doi:10.1037/h0093759

Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends in Cognitive Sciences*, *19*(7), 394–405. doi:10.1016/j.tics.2015.05.004

Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., & Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, *78*(2), 364–375. doi:10.1016/j.neuron.2013.01.039

Sugase-Miyamoto, Y., Liu, Z., Wiener, M. C., Optican, L. M., & Richmond, B. J. (2008). Short-term memory trace in rapidly adapting synapses of inferior temporal cortex. *PLoS Computational Biology*, *4*(5), e1000073. doi:10.1371/journal.pcbi.1000073

Supèr, H., Spekreijse, H., & Lamme, V. A. (2001). A neural correlate of working memory in the monkey primary visual cortex. *Science*, *293*(5527), 120–124. doi:10.1126/science.1060496

Trapp, S., & Lepsien, J. (2012). Attentional orienting to mnemonic representations: Reduction of load-sensitive maintenance-related activity in the intraparietal sulcus. *Neuropsychologia*, *50*(12), 2805–2811. doi:10.1016/j.neuropsychologia.2012.08.003

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, *41*(1), 67–85. doi:10.1016/0001-6918(77)90012-9

Williams, M., Hong, S. W., Kang, M.-S., Carlisle, N. B., & Woodman, G. F. (2013). The benefit of forgetting. *Psychonomic Bulletin & Review*, *20*(2), 348–355. doi:10.3758/s13423-012-0354-3

Appendix

Zusammenfassung
(German summary)

Menschliches Verhalten basiert zu einem großen Teil auf visueller Kognition, der Verarbeitung visueller Informationen über externe Objekte. Zwei Funktionen visueller Kognition scheinen besonders wichtig für zielgerichtetes Verhalten. Als erste Funktion ermöglicht es die Objekterkennung, Objekte in der Umwelt als Mitglieder von Objektkategorien zu identifizieren, so dass sie zur Erfüllung von Aufgaben genutzt werden können. Als zweite Funktion ermöglicht es die Kurzzeiterkennung, zu erkennen ob ein Objekt in der Umgebung kürzlich gesehen wurde. Dadurch erst kann aktuelles Verhalten bezüglich des Objekts mit früherem Verhalten in Beziehung gesetzt werden. Beide Funktionen teilen eine Einschränkung: Sie müssen über distinke Episode visueller Verarbeitung hinweg erfüllt werden, die durch Änderungen der Vearbeitungsanforderungen unterbrochen werden.

Für die Objekterkennung bergen visuelle Verarbeitungsepisoden das Problem der *selektiven Integration*. Es muss entschieden werden, ob Objektinformationen der aktuellen Verarbeitungsepisode die der Vorherigen erneuern und so mit diesen integriert werden sollen. Alternativ ist auch die getrennte Repräsentation der Objekte beider Episoden möglich. Die Entscheidung ist bedeutend. Erneuerung und Integration sollte eine kumulative und schnelle Objekterkennung ermöglichen. Objektveränderungen sollten durch Integration jedoch verdeckt werden, da deren Wahrnehmung einen Vergleich zweier Repräsentationen erfordert, hier jedoch nur eine Repräsentation vorliegt. Eine Trennung von Objektrepräsentationen birgt zwar Vorteile für die Veränderungswahrnehmung, sollte die Erkennung von Objekten jedoch beeinträchtigen. Der Grund dafür ist, dass die Objekterkennung mittels begrenzten Verarbeitungsressourcen geschieht, die im Falle getrennter Objektrepräsentationen auf diese verteilt werden müssten.

Für die Kurzzeiterkennung führen visuelle Verarbeitungsepisoden zum einem *Abgleichsproblem*. Das heißt, dass ein Objekt der aktuellen Episode (unabhängig von dessen Objektkategorie) mit Repräsentationen von Objekten abgeglichen werden muss, die nicht nur der vorherigen, sondern mehrerer kürzlich zurückliegenden Episoden entstammen können.

Ziel der vorliegenden Dissertation ist es, einen ersten Schritt zum Verständnis der Mechanismen zu gehen, die die Probleme lösen, die sich aus visuellen Verabeitungsepisoden ergeben und somit Objekt- und Kurzzeiterkennung über Episoden hinweg ermöglichen. Zu diesem Zweck führten wir fünf empirische Studien zu Fragen durch, deren Klärung zur Entwicklung einer Theorie der Objekt- und Kurzzeiterkennung erforderlich ist.

Die ersten drei Studien konzentrierten sich auf Objekterkennung, die über die visuellen Verarbeitungsepisoden von Fixationen der Augen hinweg stattfindet. Fixationen bezeichnen Perioden der visuellen Informationsaufnahme, in denen die Augen relativ stillstehen. Sie werden unterbrochen durch schnelle sakkadische Augenbewegungen. Sakkaden sind zur Objekterkennung nötig, weil sie die Fovea im Zentrum der Retina des Auges auf interessierende Objekte richten, damit diese scharf gesehen werden. Bilder von Objekten auf der Retina werden duch Sakkaden verschoben und verändert und die Aufnahme visueller Informationen wird durch sie unterbrochen. Daher teilen Sakkaden die visuelle Verarbeitung in Episoden distinkter Fixationen, denen die Mechanismen zur Objekterkennung begegnen müssen.

In zwei Studien (Poth et al., 2015; Poth & Schneider, 2016a) untersuchten wir, wie das Problem der selektiven Integration gelöst wird, um die Objekterkennung über aufeinanderfolgende Fixationen hinweg zu unterstützen. Genauer untersuchten wir die aktuelle Hypothese, dass das Problem durch einen Mechanismus gelöst wird, der ein Objekt vor und nach der Sakkade auf Korrespondenz (bzw. Kontinuität) hin testet (Schneider, 2013). Wird Objektkorrespondenz festgestellt, sollten das prä- und postsakkadische Objekt in einer gemeinsamen Repräsentation integriert werden. Wird Objektkorrespondenz hingegen gebrochen, dann sollten das prä- und postsakkadische Objekt getrennt repräsentiert werden. Diese Trennung sollte Vergleiche der Objekte ermöglichen und so die Diskrimination von transsakkadsichen Objektverschiebungen verbessern. Die Objekterkennung sollte jedoch unter der Trennung leiden, weil dazu nötige begrenzte Verarbeitungsresourcen auf mehrere statt eine Repräsentation aufgeteilt werden müssen. Die Ergebnisse beider Studien stützten diese Hypothesen. Wurde die Objektkorrespondenz nach einer Sakkade zum Objekt durch kurzzeitiges Auslassen des Objekts gebrochen, verbesserte sich die Diskrimination von Objektverschiebungen, aber verschlechterte sich die Objekterkennung. Dies zeigt, dass Objektkorrespondenz über die Sakkade sich auf die Objekterkennung nach der Sakkade auswirkt. Weitere Experimente untersuchten die Natur der Objektkorrespondenz. Sie zeigten, dass eine Beeinträchtigung der Objekterkennung auch dann auftritt, wenn die Objektkorrespondenz durch transsakkadische Veränderungen der Kontrastpolarität (und Luminanz) eines Objekts, dessen Farbe-und-Luminanz oder dessen Farbe allein gebrochen wurde. Zusammen mit dem erstgenannten Befund bedeutet dies, dass die Objektkorrespondenz über die Sakkade sowohl auf den raumzeitlichen als auch auf den Oberflächenmerkmalen eines Objekts basiert.

In der dritten Studie (Poth & Schneider, 2016b, eingereicht) untersuchten wir nun die Grenzen der Objekterkennung über Sakkaden. Da Objekterkennung begrenzte Verarbeitungsresourcen erfordert, kann sie zu einer Zeit nur für wenige Objekte erfolgen. Hier untersuchten wir, ob verschiedene Objekte über Sakkaden hinweg um diese begrenzten Resourcen konkurrieren müssen. In diesem Fall verlangsame sich die visuelle Verarbeitung nach der Sakkade mit steigender Anzahl gesehener Objekte vor der Sakkade. Die Ergebnisse der Studie zeigen, dass Objekte über die Sakkade um Verarbeitungsressourcen konkurrieren und so die Objekterkennung verlangsamen. Dies geschieht jedoch nur, wenn die Objekte aufgabenrelevant sind. Diese Ergebnisse stützen die Kernvorhersage einer aktuellen Theorie, nämlich, dass die Bedeutsamkeit einer Objektrepräsentation darüber entscheidet, ob sie die Sakkade überdauert und anschließend Verarbeitungsressourcen verbraucht (Schneider, 2013).

Mit der vierten Studie (Poth & Schneider, 2016c) wandten wir uns nun von den Verarbeitungsepisoden aufeinanderfolgender Fixationen denen zu, die durch erscheinende und verschwindende Objekte und betreffende Aufgabenanforderungen definiert sind. Hier fragten wir nach der Beziehung zwischen den Mechanismen, die die Objekt- und die Kurzzeiterkennung leisten. Es wird angenommen, dass die visuelle Verarbeitung zur Objekterkennung mit der Enkodierung in ein kapazitätsbegrenztes visuelles Arbeitsgedächtnis endet, durch das Objekte zum Bericht zur Verfügung stehen. Wir untersuchten, ob die Enkodierung ins visuelle Arbeitsgedächtnis nun nicht nur zur Objekterkennung in dieser sondern auch zur Kurzzeiterkennung in späteren Verarbeitungsepisoden nötig ist. Die Ergebnisse sprachen dafür, indem sie zeigten dass Objekte, die vermutlich nicht ins Arbeitsgedächtnis gelangten, später nicht zur Kurzzeiterkennung genutzt werden konnten. Dies bedeutet, dass die anfänglichen Schritte visueller Verarbeitung, die vor der Enkodierung ins Arbeitsgedächtnis ablaufen zur späteren Kurzzeiterkennung nicht ausreichen. Das visuelle Arbeitsgedächtnis scheint daher zur Lösung des Abgleichproblems beizutragen, indem es die Informationsmenge limitiert, die bei der Kurzzeiterkennung berücksichtigt wird.

In der fünften Studie (Poth & Schneider, 2016d, eingereicht) setzten wir die Untersuchung der Kurzzeiterkennung fort, indem wir fragten wie sie in einer vorausgegangenen Verarbeitungsepisode vorbereitet werden kann. Wir testeten, wie sich die Priorisierung von Objektrepräsentationen im visuellen Arbeitsgedächtnis auf zwei distinkte Leistungskomponenten einer bevorstehenden Kurzzeiterkennungsaufgabe auswirkt. Unsere Ergebnisse zeigten, dass eine solche Priorisierung die Gedächtnisleistung steigert, jedoch auch die visuelle Verarbeitung von Objekten zur Kurzzeiterkennung in einer zukünftigen Episode beschleunigt. Dies zeigt, dass Änderungen der Verarbeitungsprioritäten zu momentan ablaufenden Lösungen des Abgleichproblems der Kurzzeiterkennung beitragen.

Zusammengenommen, zeigen die fünf Studien wie die Mechanismen der Objekt- und Kurzzeiterkennung bestimmte Probleme bewältigen, die sich aus der Unterteilung visueller Vearbeitung in distinkte Episoden ergeben. In diesem Sinne weisen die Studien auf visuelle Verarbeitungsepisoden als eine Problemquelle für die Objekt- und Kurzzeiterkennung hin, die in aktueller Forschung größtenteils vernachlässigt wird. Umgekehrt, geben die Studien jedoch auch Anlass, über den funktionalen Wert visueller Verarbeitungsepisoden zu spekulieren.

Erklärung der selbstständigen Abfassung
(Declaration of dissertation authorship)

Hiermit erkläre ich, dass ich die vorliegende Dissertation eigenständig, ohne die Hilfe Dritter, verfasst habe und dabei keine anderen als die angegebenen Hilfsmittel benutzt habe. Alle Textstellen, die wörtlich oder sinngemäß anderen Quellen entstammen, habe ich als solche kenntlich gemacht. Die Dissertation hat in der gegenwärtigen oder einer anderen Fassung keiner anderen Fakultät oder Universität vorgelegen.

Christian Hendrik Poth

16.01.2017

Bielefeld, Datum, Unterschrift

Urheberschaftserklärung
(Declaration of manuscript authorship)

Hiermit erklären wir, dass alle nachfolgend aufgelisteten Manuskripte eigenständig von Herrn Poth verfasst wurden. Herr Poth war hauptverantwortlich für die Versuchsplanung, Untersuchungsdurchführung und Datenauswertung der empirischen Studien. Darüber hinaus verfasste und überarbeitete Herr Poth alle Manuskripte selbstständig. Die gesamte empirische Arbeit und die Verfassung aller Manuskripte entstand unter Betreuung von und in Ko-Autorenschaft mit Herrn Professor Dr. Werner Schneider. Des Weiteren wirkte Herr PD Dr. Arvid Herwig als Ko-Autor an der ersten Studie und des entsprechenden Manuskripts mit.

Prof. Dr. Werner Schneider   PD Dr. Arvid Herwig   Christian Hendrik Poth
                                    (für Studie 1)

16.1.2007   *16.01.2017*   *16.01.2017*

Bielefeld, Datum, Unterschrift   *Bielefeld, Datum, Unterschrift*   *Bielefeld, Datum, Unterschrift*

Studie 1 Poth, C. H., Herwig, A., & Schneider, W. X. (2015). Breaking object correspondence across saccadic eye movements deteriorates object recognition. *Frontiers in Systems Neuroscience, 9*, 176. doi: 10.3389/fnsys.2015.00176

Studie 2 Poth, C. H., & Schneider, W. X. (2016). Breaking object correspondence across saccades impairs object recognition: The role of color and luminance. *Journal of Vision, 16*(11), 1. doi: 10.1167/16.11.1

Studie 3 Poth, C. H., & Schneider, W. X. (2016). Attentional competition across saccadic eye movements. *Submitted for publication in the Journal of Experimental Psychology: Human Perception and Performance*.

Studie 4 Poth, C. H., & Schneider, W. X. (2016). Episodic short-term recognition requires encoding into visual working memory: Evidence from probe recognition after letter report. *Frontiers in Psychology, 7*, 1440. doi: 10.3389/fpsyg.2016.01440

Studie 5 Poth, C. H., & Schneider, W. X. (2016). Prioritization in visual working memory enhances memory-retention and speeds up processing in a comparison task. *Submitted for publication in Attention, Perception, and Psychophysics*.

Bestätigung der Einreichung von Manuskripten zur Publikation
(Confirmation of manuscript submissions for publication)

Betreff: Submission Confirmation for Attentional competition across saccadic eye movements - ███████████████████████

Von: "Journal of Experimental Psychology: Human Perception and Performance" <em@editorialmanager.com>

Datum: 09.12.2016 18:25

An: "Christian H. Poth" <c.poth@uni-bielefeld.de>

Dear Mr. Poth,

Your submission "Attentional competition across saccadic eye movements" has been received by Journal of Experimental Psychology: Human Perception and Performance.

You will be able to check on the progress of your submission by logging on to Editorial Manager as an author. The URL is http://xhp.edmgr.com/.

Your manuscript will be given a reference number once an Editor has been assigned.

Please note that you may also confirm or Authenticate your ORCID iD by clicking here http://xhp.edmgr.com/l.asp?i=84842&l=F7TZ3FCO.

Best regards,
Editorial Office
Journal of Experimental Psychology: Human Perception and Performance

APA asks that you please take a moment to give us your feedback on the submission process, by completing a short survey, available at ████████████████████████████

Betreff: Attention, Perception, & Psychophysics - Manuscript ███████████
Von: "Attention, Perception, & Psychophysics "
<onbehalfof+journals+psychonomic.org@manuscriptcentral.com>
Datum: 08.12.2016 18:28
An: c.poth@uni-bielefeld.de

08-Dec-2016

Dear Mr. Poth:

Your manuscript entitled "Prioritization in visual working memory enhances memory-
retention and speeds up processing in a comparison task" has been successfully
submitted online and is presently being given full consideration for publication in
the Attention, Perception, & Psychophysics.

Your manuscript ID is ████████████

Please mention the above manuscript ID in all future correspondence or when calling
the office for questions. If there are any changes in your street address or e-mail
address, please log in to Manuscript Central at https://mc.manuscriptcentral.com
/pandp and edit your user information as appropriate.

You can also view the status of your manuscript at any time by checking your Author
Center after logging in to https://mc.manuscriptcentral.com/pandp.

Thank you for submitting your manuscript to the Attention, Perception, &
Psychophysics.

Sincerely,
Attention, Perception, & Psychophysics Editorial Office