Behavioral modulation in a comparative visual search task

An in-depth analysis of chunk and revisit fixations

Masterarbeit

der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen

vorgelegt von

Linda Carmen Bräutigam

Tübingen, Januar 2021

I. Erklärung

Hiermit erkläre ich, Linda Bräutigam,

- dass ich diese Arbeit selbst verfasst habe.
- dass ich keine anderen als die angegebenen Quellen benutzt und dass ich alle wörtlich oder sinngemäß aus anderen Werken übernommenen Aussagen als solche gekennzeichnet habe.
- dass die eingereichte Arbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens gewesen ist.

Tübingen, den 25.01.2021

II. Overview

III. Abstract

Despite of capacity and temporal limitations of working memory, humans are able to solve complex tasks. This thesis opts to shed light on the underlying mechanisms and behaviors that counteract those limitations.

The first part of this thesis contains a theoretical analysis. First of all, the topic working memory was examined focusing on its limitations and basic mechanisms. Those limitations can be counteracted by mechanisms like rehearsal, combining information and filtering. An in-depth analysis of the associated terms grouping, clustering and chunking revealed a general interchangeability of those three terms as well as slight differences. Lastly the topic of eye movements with the example of fixation duration modulation as well as advantages of eye tracking in comparison to behavioral data were reviewed. Besides other factors discussed, eye tracking data showed to be advantageous due to its inclusion of spatiotemporal aspects.

The second part consists of a systematic analysis of eye tracking data recorded in a comparative visual search paradigm. An evaluation of eye tracking in comparison to behavioral data showed a comparability of those two data formats. The mechanisms of assigning fixations to the non-object or object category was examined and showed an overall good performance. To investigate individual search behavior several parameters like fixation duration, number of fixations, total fixation time and lastly the number of hemifield switches were analyzed. The delay duration, object complexity and number of switches varied, leading to a behavioral modulation and strategy adaption in order to suffice the altered circumstances. Commencing on trial level, investigations were also conducted on a switch level as an indicator of a cognitive processing step. Using the level of a switch enabled a differentiation of fixations between *chunk* and *revisit.* This resulted in new conclusions proposing chunk size as a valid indicator of working memory capacity, the use of revisits as a consolidation of information in working memory and differences regarding those two fixation types in relation to the task. Using this fixation type classification, several parameters were inspected, resulting in new insights about processing mechanisms and behavioral modulation. In the final step, correlations between several parameters were analyzed, including the operationalizations of the acquisition and memorization strategy. Chunk size thereby, showed to be a better approximation for working memory strategy use than the previously used processing time.

IV. Zusammenfassung

Trotz der Arbeitsgedächtnislimitationen hinsichtlich der Kapazität und zeitlichen Dauer von Information sind wir in der Lage herausfordernde Aufgaben ohne größere Probleme zu lösen. Diese Arbeit untersucht die zugrundeliegenden Mechanismen und Verhaltensweisen, die dies ermöglichen.

Der erste Teil dieser Arbeit befasst sich mit den theoretischen Hintergründen des Arbeitsgedächtnisses. Dabei werden dessen Limitationen und grundsätzliche Mechanismen näher erklärt. Diesen Limitationen kann durch Aufrechterhaltungsprozesse (engl. „*rehearsal*“) sowie dem Kombinieren und Filtern von Information entgegengewirkt werden. Eine ausführliche Analyse der assoziierten Begriffe: „*chunking*”, „*clustering*” und „*grouping*”, welche sich auf das Gruppieren von Informationen beziehen, zeigte eine große Ähnlichkeit dieser Begriffe sowie einige Unterschiede. Zuletzt wird auf die Bedeutung von Augenbewegungen am Beispiel der Modulation der Fixationsdauer eingegangen und die Vorteile von Augenbewegungsdaten im Vergleich zu behavioralen Daten aufgezeigt. Vor allem die räumlich-zeitliche Datenstruktur zeichnet sich neben weiteren besprochenen Aspekten als größter Vorteil von Augenbewegungsdaten ab.

Der zweite Teil besteht aus einer empirischen Analyse von Augenbewegungsdaten einer vergleichenden visuellen Suchaufgabe. Eine Evaluierung von Augenbewegungsdaten und der Vergleich mit behavioralen Daten zeigte eine Vergleichbarkeit der beiden Datenformate. Die Zuordnung von Fixationen zu einem oder keinem Objekt wurde genauer untersucht und brachte neue Erkenntnisse. Um das individuelle Suchverhalten zu untersuchen, wurden mehrere Parameter wie Fixationsdauer, Anzahl an Fixationen, gesamte Fixationszeit sowie Anzahl der Hemifeld-Wechsel berücksichtig. Die Variationen der Verzögerung zwischen den Maskenwechseln (engl. „*delay*“), der Komplexität der Objekte sowie der Anzahl der verwendbaren Seitenwechsel führten zu einer Verhaltensmodulation und Strategieanpassung der Versuchsteilnehmer. Analysen wurden sowohl auf der Ebene eines Trials, als auch auf der Ebene eines Seitenwechsels durchgeführt, wobei ein Seitenwechsel als Indikator für einen kognitiven Verarbeitungsschritt genutzt wurde. Letztere Ebene ermöglichte eine Differenzierung von Fixationen in erstmalige Fixation eines Objektes (engl. „*chunk*“) und in Refixationen (engl. „*revisits*“). Dies liefert neue Erkenntnisse welche die Anzahl von *chunk* Fixationen als validen Indikator für Arbeitsgedächtniskapazität nahelegen. Zusätzlich konnte gezeigt werden, dass Refixationen zur Konsolidierung von Information im Arbeitsgedächtnis genutzt werden und sich Dauer und Anzahl abhängig von der Schwierigkeit der Aufgabe unterscheiden. In einem letzten Schritt wurden Korrelationen zwischen den zuvor genannten Parameter untersucht, mitunter auch die Operationalisierung von Akquisitions- und Memorisierungsstrategie. Dabei wurde festgestellt, dass *chunk size* eine bessere Approximation der Memorisierungsstrategie darstellt als die in der Literatur verwendete Prozessierungszeit.

V. Theoretical analysis

1 Working memory

The concept of working memory based on short term memory was first introduced by Miller et al. (1960) and further established by Atkinson and Shiffrin (1968) and Baddeley and Hitch (1974).It encompasses the temporary storage of small amounts of information. As proposed by Atkinson and Shiffrin (1968) working memory is a part of short term memory which controls information flow into and out of long-term memory. Therefore, working memory is known to be crucial in cognition and for mental tasks like problem solving and planning. However, because there were studies examining patients with short term memory deficits, that had no impairments Baddeley and Hitch (1974) proposed a three-component model which tried to suffice for those findings. In their model a central executive acting as an attentional control system, is aided by two short-term storage systems: the visuo-spatial sketchpad and the phonological loop. Those two parts process visual and verbal acoustic material respectively. Later on, an episodic buffer was added in this now multidimensional model (Baddeley, 2000). This fourth component is limited in its capacity and can hold around four multidimensional episodes or chunks and provides a temporal store. The episodic buffer is capable of binding together information from different modalities like acoustic and visual information (Baddeley, 2003). First this store was thought to have an active and attentionally demanding role, as it could bind together different information (Baddeley, 2007), which later changed to a passive storage (Baddeley et al., 2009). A newer overview of this approach can be found in Baddeley (2010).

An alternative definition for working memory stems from Cowan (1999). According to him working memory refers to a cognitive process, which retains information in an accessible state. Working memory is the collectivity of mnemonic functions that preserve information. His embedded-processes model of working memory is based on five principles emphasizing the link between memory and attention, which are mentioned in the following.

First, information in the working memory comes from three hierarchical faculties: inactive but retrievable long-term memory, active long-term memory without the focus of attention and activated memory with the focus of attention and awareness. Each faculty possesses different processing limits. While the focus of attention is limited in its capacity, the activation stage has temporal limitations. Third, the focus of attention is controlled by voluntary as well as involuntary processes, namely the *central executive* and the attentional orienting system respectively. Even though there in no awareness of some unchanged and unimportant features of stimuli they can still activate connected features in memory. Lastly awareness influences information processing for example in the number of encoded features in perception. For a in depth assessment of this approach see Cowan (2016).

1.1 Working memory limitations

Working memory is known to be limited in time and capacity. Those limitations are now elaborated further.

Regarding limitations in time early studies from Brown (1958) and Peterson and Peterson (1959) showed a temporal decay of information in the short-term memory within seconds if no rehearsal took place. According to Cowan (1999) working memory is limited in the activation time, which fades within 10 to 20 s unless reactivated. Temporal limitation of working memory is confirmed (Barrouillet et al., 2011, Barrouillet and Camos, 2012, Barrouillet et al., 2012), rejected (Lewandowsky and Oberauer, 2009, Lewandowsky et al., 2009) and discussed by several sources (Cowan, 2008, Portrat et al., 2008, Oberauer et al., 2016). Even though temporal limitations of working memory might be in debate, because working memory is thought to be a subpart of short-term memory in this thesis a temporal limitation of working memory is assumed.

The capacity limitation of working memory on the other hand was investigated much more thoroughly and less controversial. Presumably the most famous capacity limit for working memory was set by Miller (1956a) as the magical number seven plus minus two meaningful items or chunks. Cowan suggested in several of his studies that the focus of attention, which could be seen as the limitation of working memory capacity, is limited to three to five items or chunks (Cowan, 1992, Cowan, 1999, Cowan, 2001). Another approach of working memory limitations, namely working memory span, is also linked to general measures of cognition and is usually around three to four chunks (Baddeley, 1992). This measure is useful in the prediction of a range of cognitive tasks like logical reasoning. It is to note, as Engle and Kane (2004) pointed out, short-term-memory span itself and working memory capacity seem to differ from another. While working memory capacity or presumably also working memory span are good predictors of cognitive task performance, short-term-memory span is not.

This capacity limitation can however be enhanced by chunking and structuring of information (Cowan, 1999). Cognitive control like suppressing irrelevant information (Engle and Kane, 2004), the ability to switch or divide attention (Barrouillet et al., 2004) and efficient updating and maintenance of information (Miyake et al., 2000) seems to also have an important role in working memory use itself.

Such a capacity limit is hard to assess because it varies vastly across subjects, lifespan and tasks (Cowan, 2010). It is still not fully understood which mechanisms help to counteract the limited working memory capacity (Cowan, 2010), but some of them are examined in chapter 2: Counteracting working memory constraints. Furthermore, it may not be only a storage limitation but also a difference in effectiveness of working memory use (Kane et al., 2001).

1.2 Basic mechanisms of working memory

According to Cowan (1999) there are four basic mechanisms of working memory: encoding, representation, maintenance and retrieval of information.

A stimulus is encoded in working memory when stimulus feature activate memory. This activation gets better when attention is allocated to the item to be encoded (Cowan, 2010). When a stimulus is encoded in working memory an internal representation of the items is formed in which properties can be grouped together to form a coherent representation of the item in question (Cowan, 2010). After a representation is formed it has to be maintained in working memory or else it is lost. Usually maintenance is done be reactivation of information for example by verbal rehearsal as suggested by (Baddeley, 1992). Alternatively maintenance can be achieved by focusing attention on specific aspects of the stimulus (Cowan, 1999). By searching through a set of items reactivation can also occur (Cowan, 1992). After some time, stored information has to be retrieved out of memory into the focus of attention. The retrieval from active memory and long-term memory differs in the time needed. While retrieval from active memory is rather quick, retrieval from long-term memory is much slower (Cowan, 1999).

Similar to Cowan (1999), Braisby and Gellatly (2012) suggested a three stage model of memory. According to their model memory consists of an encoding, a storage, and a retrieval stage. In contrast to Cowan (1999) the representation mechanism in this three-stage model is encompassed in the encoding stage. Storage is thereby similar to maintenance in Cowans' working memory model coming more into play when looking at non-active memory. In the retrieval stage, information is reactivated for direct use (Braisby and Gellatly, 2012). Even though the model from Braisby and Gellatly (2012) is more general and not designed specifically for working memory there still are some parallels.

There is a newer proposal from Eriksson et al. (2015) relating to this topic. He stated that the basic feature of working memory is the short-term maintenance of no longer available sensory information. Information maintenance is the result of an interaction between a selective attention process and related long-term memory representations. Similar to Cowan (1999) and Braisby and Gellatly (2012), the working memory model suggested by Eriksson et al. (2015) focuses on the maintenance and retrieval of information. When large amounts of information do not "fit" in the focus of attention additional rehearsal processes are needed to maintain information. In their work the retrieval phase is seen as a delayed-match-to-sample task in which selective attention and pattern completion processes are used to (re)form information (Eriksson et al., 2015). This component process view of working memory aligns with the two previously discussed models.

2 Counteracting working memory constraints

The previously discussed limitations of working memory raise the question why we usually do not notice such a limitation in our ability to perform high cognitive tasks. Besides the obvious answer that we are used to these limitations, there are some processes involved in helping us to overcome these limitations.

Engle and Kane (2004) stated that working memory use can be maximized by using coding strategies, grouping and procedures for maintaining activation. Such processes could be phonological, visual, spatial, motoric, or auditory. In their two-factor model Engle and Kane (2004) explain how individual differences in working memory capacity or executive attention can lead to performance differences. They furthermore stated the importance of executive attention in maintaining information and in resolving competition between task-appropriate and task-inappropriate responses. To produce this model Engle and Kane (2004) considered seven alternative hypothesis why some people perform better in different working memory tasks, which were all found to be insufficient by themself. Those hypotheses ranged from a difference in mental effort or motivation over a speed hypothesis to a rehearsal-difference-hypothesis, strategy allocation hypothesis, general processing hypothesis, task specific hypothesis and even considered differences in word knowledge.

Alongside verbal or nonverbal rehearsal, grouping or chunking (Allon et al., 2019) as well as the ability to filter information (Kane and Engle, 2003) are the main strategies of how to increase working memory capacity and maximize information maintenance. Those thee strategies will be discussed in detail below.

2.1 Rehearsal

Adopting the focus of attention theory Eriksson et al. (2015) stated that information can be maintained though reverberating signals between different brain regions, as long as the content of information does not exceed the focus of attention. If too much information is needed to be maintained, additional rehearsal processes are necessary to complement the active maintenance process of information. This neural process could be seen as some sort of inherent and automatic rehearsal process.

Now for the active rehearsal processes. Similar to the distinction between spatial and verbal working memory from Baddeley and Hitch (1974) active, attention-based rehearsal processes can also be differentiated between spatial and verbal rehearsal.

First spatial rehearsal will be considered. When looking at spatial rehearsal the oculomotor rehearsal hypothesis should be discussed. First proposed by Baddeley (1986) the hypothesis suggests that spatial working memory is mediated by implicit eye movements, later called oculomotor preparation (Pearson et al., 2014). Spatial information is connected to gestalt properties like color and shape and be planning saccades or moving the eyes to the position of a certain object this information is reactivated (Baddeley, 1986, Pearson et al., 2014) . Based on three experiments Awh et al. (1998) proposed that the maintenance of spatial information is mediated by shifts of spatial attention itself and not necessarily eye movements. After assessing the influence of spatial locations on working memory accuracy they discussed how the oculomotor rehearsal hypothesis could fit their results. Pearson et al. (2014) used a version of the abducted-eye paradigm to tackle the question whether oculomotor rehearsal or covert spatial attention shifts could account for the increased performance in spatial working memory tasks. In their task it was either physically impossible to perform saccadic eye movements to target positions by rotating the participants head to an abducted position (abducted-eye condition) or it was possible to perform eye movements due to a frontal head position (normal condition). They found that eye-abduction only impaired working memory performance when potential eye movements would have been impossible to make due to the abduction of the eyes. Therefore, they concluded that spatial working memory performance depends on eye movement rehearsal and not just spatial attention. Similar findings supporting the spatial rehearsal hypothesis using (planned) eye movements were obtained by Awh et al. (1999), Tremblay et al. (2006), Smyth (1996), and Pearson et al. (2014).

The second rehearsal process that is discussed is phonological or also called articulatory rehearsal. A maintenance of information in verbal working memory through articulatory rehearsal was first proposed by Baddeley (1986). First called articulatory rehearsal (Baddeley et al., 1984, Longoni et al., 1993) this process was later renamed phonological rehearsal, due to the rename of articulatory loop to phonological loop (Baddeley, 1986, Baddeley et al., 1984). This rehearsal process can be differentiated between verbal also called overt articulatory and non-verbal also called covert articulatory rehearsal. As these names suggest verbal or non-verbal rehearsal relate to a speech based maintaining process of information. Many early studies investigating the effect of verbal and nonverbal rehearsal were done in children (Daehler et al., 1969, Flavell et al., 1966, Keeney et al., 1967, Torgesen and Goldman, 1977, Kirchner and Klatzky, 1985). An interesting study from Hourihan et al. (2009) indicates that verbal rehearsal results in better encoding compared to nonverbal rehearsal. In two experiments using the *item method directed forgetting* paradigm (For more details on this paradigm see Basden et al., 1993) they compared memory performance in conditions were verbal rehearsal was allowed and prevented. In conditions were verbal rehearsal was prevented by a verbal suppression task, nonverbal rehearsal occurred and participants showed an enhanced memory performance in comparison to the no rehearsal condition (Hourihan et al., 2009). Further neurological evidence for verbal rehearsal and its positive effects on memory maintenance was provided by Awh et al. (1996) using positron emission tomography and Hwang et al. (2005) using EEG recordings.

2.2 Combining information

Since working memory capacity is limited, one can optimize needed resources by grouping also called chunking items together into integrated units of information (Treisman, 1982). The fact that grouping affects learning and memory is known since nearly 80 years (Katona, 1940).

In a review from Wagemans et al. (2012) they examined perceptual grouping. Defined as some elements of a visual field fit together better than others, perceptual grouping is a form of perceptual organization (Wagemans et al., 2012). First introduced by Wertheimer (1923), the term perceptual grouping has a long history, which will not be discussed here. In general items can be grouped in a variety of ways for example by: spatial and temporal proximity, similarity for example in color, size or orientation, common fate, direction, symmetry parallelism, continuity or closure as well as common regions to name a few (Wagemans et al., 2012).

Because there are so many features that can be used, only some of them will be examined in the following in more detail: first grouping by spatial proximity, followed by grouping by gestalt features like color, shape or orientation.

The factor of relative distance, called (spatial) proximity was first principle of grouping introduced by Wertheimer (1923). In his experiment he showed that equally spaced dots were not grouped together but by altering the distance participants grouped those dots together. In a new study from Allon et al. (2019) participants had to perform a change-detection task under different conditions. It could be shown that in conditions were target items or distractors were grouped as a triangle, indicating proximity, participants performed better. It could be shown several times that spatial proximity will lead to grouping (Kubovy et al., 1998, Rock and Brosgole, 1964) and subsequently increases memory precision (Allen and Haun, 2004, Theeuwes, 1996, De Lillo, 2004).

Quinlan and Cohen (2012) used a change detection paradigm and *single-probed recognition* to investigate grouping. In their experiment they found that there was a strong color shared effect, as memory was better for items that had the same color. To group items spatial proximity or color can be used as valid features, shape on the other hand is a less used cue for grouping (Quinlan and Cohen, 2012). Another well-known example for perceptual grouping was conducted by Treisman (1982). In four *conjunction search* experiments she used color and shapes to investigate the effects of grouping. Her results confirmed the previously formed feature map theory, in which as the name suggests features are coded in distinct feature maps (Treisman and Gelade, 1980). Later evidence reinforced the suggestion that features are coded in different maps so grouping will likely take place within one feature category (Chen et al., 2020). There are many more examples that show the beneficial effects of grouping on perceptual processing (Baylis and Driver, 1992, Kimchi et al., 2007, Kimchi et al., 2016, Wagemans et al., 2012, Li et al., 2018b).

One of the best known statements regarding the use of grouping in the optimal use of memory was done by Miller (1956b). In the context of grouping he introduced the word “chunk”, which many other adopted (Mandler, 1967, Egan and Schwartz, 1979). An in-depth analysis of what can be understood under the term chunking can be found in the chapter 3: Grouping, clustering and chunking.

2.3 Filtering

As attention can be allocated voluntarily though a top down process or involuntarily due to high salient stimuli, it is suggested that the individual ability to filter out distractors could predict differences in working memory performance (Allon et al., 2019).

Using the *Stroop task* Kane and Engle (2003) were able to confirm the importance of executive control and goal maintenance on selective attention and therefore also working memory performance. By using this paradigm, the goal to ignore or filter out the written word was directly included.

Another well-known study connecting the ability to filter out unimportant information to working memory performance was conducted by Vogel et al. (2005). They observed that the ability to select important information was highly predictive of working memory capacity. Vogel et al. (2005) furthermore suggested that the difference between *high* and *low capacity* individuals might result not from the capacity itself but from the ability to allocate attention or rather working memory capacity. These finding was affirmed by McNab and Klingberg (2008) with further neural data using activity changes associated with filtering.

In addition to neural data there are several behavioral approaches indicating the same positive effect of filtering on working memory related performance (Fukuda and Vogel, 2009, Allon and Luria, 2017, Arnell and Stubitz, 2010, Li et al., 2017).

The huge body of evidence regarding the importance of filtering information and controlling working memory capacity allocation uses the distinction between high and low capacity individuals (Fukuda and Vogel, 2009, Allon and Luria, 2017, Arnell and Stubitz, 2010, Li et al., 2017, Vogel et al., 2005, Cowan and Morey, 2006, Jost et al., 2011). Despite this distinction indicating a difference in capacity most studies suggest that the identified (working) memory capacity difference is the result of the ability to successfully ignore distractors and focus on the task instead of a real difference in capacity. These studies furthermore indicate that filtering mechanisms act as a gateway to select task relevant information. There is a consensus that filtering mechanisms are individual capabilities to selectively encode and maintain relevant information in the presence of irrelevant information (Robison et al., 2018).

3 Grouping, clustering and chunking

As investigated before, combining features or objects is a useful mechanism to counteract working memory limitations. Even though this mechanism is well studied, there is no unified nomenclature regarding this subject. Three main terms can be identified: grouping, clustering, and chunking. Starting with the term chunking all three expressions will be discussed in the following.

3.1 Chunking

The term chunking is shaped by early work from the Dutch scientist de Groot (1946) in the field of problem solving and probably better known by Miller (1956b) from the field of perception and memory. Miller (1956b) defined a chunk as a unit of information that is independent of the information stored. He stated that chunk relates to the process of organizing or grouping input into units and furthermore stresses the involvement of learning in their formation. Interestingly, already Miller (1956b pp 93.) stated that “we are not very definite about what constitutes a chunk of information”, something that will continue to be a problem. One of the uses of chunking Miller (1956b) highlighted was the bypass of the informational bottleneck of working memory.

While the term chunk and the process of chunking was investigated in many different areas of research, the concept (Gobet et al., 2001) still describes a grouping procedure used to optimize a limited resource despite being diversified. In succession to Miller (1956b), Capaldi et al. (1986) defined a chunk as the grouped information of individual features into one “package” or memory unit. The process of chunking was thereby described as the process of combining separate items into larger order units (Capaldi et al., 1986). In their grouping experiments with rats they could showed that chunking is a mechanism not only present in humans. Quinlan and Cohen (2012) suggested that a chunk is a perceptual group, which can be stored in a slot in memory.

According to Gobet et al. (2001) chunking can be distinguished based on how and when chunking is assumed to occur. Goal-oriented chunking assumes a deliberate, conscious control of the process (Gobet et al., 2001). Perceptual chunking on the other hand is assumed to be more automatic and continues and occurs during perception. Similar to this proposal, there is a perceptual chunking hypothesis suggested by Simon and Gilmartin (1973), that proposes that chess masters can represent entire chess positions by a small number of chunks (Egan and Schwartz, 1979). Perceptual chunking as described by Chase and Simon (1973) and Simon and Gilmartin (1973) consists of two features: first the perception of chunks as independent and second that chunks are labeled. An alternative to the perceptual chunking hypotheses is the conceptual chunking hypothesis from Egan and Schwartz (1979). This hypothesis links chunking to concepts in long-term memory (Egan and Schwartz, 1979).

Chunking seems to be better in skilled compared to unskilled subjects and is a skill that can be learned, as identified by Egan and Schwartz (1979) in three experiments.

Even though there are several definitions for chunking they all seem to agree that chunking is a process of grouping information and storing that grouped information in memory. However, when chunking is assumed to occur as well as how this mechanism is mediated is still to be discussed.

3.2 Clustering

When talking about the term clustering Pomplun et al. (2001) should be mentioned, as they had a great influence on shaping this term. They proposed three possible search strategies that can be outlined within a *comparative visual search task* (an in depth dissection of this paradigm can be found in chapter 1.1.1: Comparative visual search of the experiment part), one of them being the clustering strategy. Interestingly they stated that those strategies are characterized by the way in which the visual scene is divided into chunks (Pomplun et al., 2001). By using this strategy, a visual scene is processed in so called object clusters, groups of objects with a similarity or homogeneity. Clusters are thereby defined by object density and heterogeneity of objects (Pomplun et al., 2001). In such a “clustering” strategy different features can be used to conjoin several objects. According to Pomplun et al. (2001) clusters can be defined either by spatial proximity or by shared object features like color. So, despite the connotation of cluster having a spatial context the cluster strategy is not restricted in this manner. Similar to the definition of chunking, Pomplun et al. (2001) stated that this strategy and thereby clustering itself could be seen as an attempt to exploit limited working memory capacity by grouping objects in a cost efficient way.

According to Xu (2010) clustering enables the formation of structures and hierarchies in visual scenes. Furthermore, it provides structure to visual displays, which might allow a better allocation of attentional resource and consequently more efficiency in searching, leading to a faster search performance (Xu, 2010). Using a feature search experiment were participants had to find a target letter among distractor letters Xu (2010) found that clustering slowed down easy feature search but speeded up difficult spatial configuration search. In comparison to Pomplun et al. (2001), Xu (2010) used the term clustering in a spatial context. Nevertheless, the term clustering was used in the connection with visual search.

Staying in the paradigm of visual search for example Cohen and Ivry (1991) used the term clustering in one of their older experiments. In this example clustering was used synonymously to grouping in the visual search paradigm. By accounting for eccentricity and item number they on the other hand found that density manipulation did not affect feature search.

It seems that the term clustering is used in the field of visual search (Sivic and Zisserman, 2006, Bruce and Tsotsos, 2009) especially when a spatial or hierarchical components are involved (Xu, 2010, Lin and Brandt, 2014). Furthermore, it seems that the term clustering is commonly used in regard to search engines and information theory (Di Giacomo et al., 2008, Cai et al., 2004, Kalantidis et al., 2016, Girod et al., 2011).

Different from Xu (2010) who stated that the use of clustered items was avoided in the beginnings of visual search research resulting in a limited understanding of clustering, I suggest that the term clustering may not have been used because it was more common to use the term chunking or simply grouping. While clustering might have more of a spatial connotation the mechanism behind seems very similar to the term chunking proposed early on.

3.3 Grouping

Other than clustering and chunking the term grouping is used as a general expression to explain the organization and combination of features based on similarities like spatial proximity or other feature similarities (Rock and Brosgole, 1964, Kubovy et al., 1998). The body of research regarding grouping in working memory is huge (Allon et al., 2019, Baylis and Driver, 1992, Capaldi et al., 1986, Chen et al., 2020, Kahneman and Henik, 1977, Kubovy et al., 1998, Li et al., 2018b, Quinlan and Cohen, 2012, Rock and Brosgole, 1964, Treisman, 1982, Wagemans et al., 2012).

It seems that many authors used grouping and either chunking (Capaldi et al., 1986), clustering (Xu, 2010) or even binding (Quinlan and Cohen, 2012) to either make it more interesting to read or because they were influenced by some other author that used the term. As discussed, those three terms are used in many different areas of research and they may have slightly different connotations. Some authors may attribute a slightly difference nuance to one of these terms or one term is more commonly used in a certain paradigm.

Nevertheless, is seems that the base principle of all those terms is very similar, and they are used to describe closely related mechanisms. This mechanism concerns the efficient structuration of items in a visual display based on similarities like spatial proximity or other object feature like color in order to exploit limited working memory capacity and use it more efficient.

4 Eye movements

During visual perception saccadic eye movements scan the environment (Findlay and Walker, 1999, Findlay et al., 2003). Due to anatomy related acuity limitations, eye movements are an essential part of human vision. Eye movements are used to overcome the limitations of the human oculomotor system (Vasilyev, 2020, Findlay and Walker, 1999, Findlay et al., 2003) and depend on cognitive processes connected to visual information processing (Trukenbrod and Engbert, 2014).

Eye movements can be differentiated between saccades and fixations. While saccades are high velocity gaze shifts used to direct vision to a new region of interest, fixations are used to maintain perceptual input at a single region of interest in a high resolution (Vasilyev, 2020). Besides fixation duration and saccade length, the number of fixations and total fixation time are often included in common temporal and spatial eye movement assessments (Manor and Gordon, 2003).

Similar to the distinction between spatial and temporal pathways in neural anatomy, models for the control of eye movements can be distinguished between those same two categories. While spatial eye movement models focus on the spatial aspects of saccadic eye movement, temporal models are used to determine the fixation duration (Trukenbrod and Engbert, 2014). Most research in this field uses reading tasks and deals with spatial aspects, while temporal models are rarely investigated.

There are high individual differences in eye movements especially in the fixation duration depending on the task, processing difficulty and health condition (Vasilyev, 2020, Van der Stigchel et al., 2013, Staub and Benatar, 2013). In the next section I will further investigate which factors can modulate fixation duration and why fixation duration differs as much as it does between individuals.

4.1 Fixation duration modulation

Fixation durations can range from less than hundred milliseconds to more than a second. Three main factors for this high variability are suggested by Trukenbrod and Engbert (2014).

First, the global properties of the used stimuli influence fixation duration. Examples for such properties are visual clutter (Henderson et al., 2009), luminance and contrast of the pictures (Loftus, 1985) as well as low- or high-pass filtering (Groner et al., 2008). All these properties can be summarized as stimulus difficulty (Gould and Dill, 1969) or complexity.

The second factor influencing fixation duration is task specific processing. While the effect of the first factor changes with stimulus properties, the effect of the second factor is modulated by the instruction independent of stimulus property (Trukenbrod and Engbert, 2014). For example memorizing a scene leads to longer fixation durations compared to searching the same scene as shown by Henderson et al. (1999). A related task specific increase of fixation duration was found in relation to memory load (Gould, 1973).

Lastly, there are large individual differences in mean fixation duration. Some of them might result from practice or age as shown by McConkie et al. (1991) and Rayner et al. (2011). Nevertheless, most differences seem to reflect individual preferences or preadaptation (Castelhano and Henderson, 2008, Rayner et al., 2007, Trukenbrod and Engbert, 2014).

Early on it was proposed by Vaughan and Graefe (1977), that fixation duration is confounded by two factors: First the cognitive processing time and secondly the time to plan the next eye movement. When cognitive processing time is examined is it necessary to decrease the factor of eye movement planning. To do so it could be advantageous to use a paradigm were objects are always placed in the same position in similar distances from another. Using this, participants can learn predefined object positions, which may decrease planning time. Furthermore, it was believed that the oculomotor latency would not determine the fixation duration in visual search tasks, that involve high information objects or many items (Moffitt, 1980). Moffitt (1978) found that fixation duration increased as a function of memory size and concluded that it was reflective of the amount of information processed. According to Moffitt, the fixation duration is determined by the amount of information revived from the number of objects per fixations, while holding the information of each item constant or by an increase in the information value of the items, when the number of objects per fixation was held constant.

Whether the control of fixation duration is involuntary as suggested by Hooge and Erkelens (1996), Hooge and Erkelens (1998), and Hooge et al. (2007) or is influenced by global and local control as suggested by Engbert et al. (2002) is to be discussed. There are a variety of computational models that opt to analyze the control of eye movements (Trukenbrod and Engbert, 2014, Najemnik and Geisler, 2005, Butko and Movellan, 2008). Vasilyev (2020) gave an overview of eye movement models and introduced a new one. In his new model Fixation duration depends on the difficulty and interacts with saccade length. He studied the ability to control Fixation duration using a visual search task. Even though this model ceased to explained differences in experimental data, it was able to reproduce the modulation depending on difficulty and included preceding and succeeding saccades.

4.2 Advantage of eye tracking compared to behavioral data

There are some advantages of eye movement measurements in comparison to behavioral data, already recognized by Pomplun et al. (2001). Early on most studies regarding visual search just used reaction times and error rates, which elicited some major findings but proved to be limited when looking at cognitive processes involved. Some major advantages were found by introducing eye tracking technology.

First of all, eye tracking data encompasses information about the duration of the search in form of fixation duration. Furthermore, the time course of the search could be reconstructed. In addition to the temporal data eye tracking also involves spatial information, so in addition to the timepoint a participant fixated an object, we would also know where the fixation was directed to. With behavioral measurements one would know how long a participant would take to solve a trial, but what participants did during a trial remains unknown. The use of spatiotemporal data can lead to more precise modeling due to exact empirical validation (Pomplun et al., 2001).

Fixations resemble the overt attention shifts and are coupled to information processing (Hoffman, 1998). Furthermore the focus of attention and can be considered a valid measure of processing time (Pomplun et al., 2001). Therefore eye tracking is a powerful tool to examine attentional processing (Franco-Watkins and Johnson, 2011). Despite in the use of decision making (Ashby et al., 2016, Gidlöf et al., 2013, Shi et al., 2013) and visual reach (Najemnik and Geisler, 2005, Kit et al., 2014, Drew et al., 2017, Zelinsky and Sheinberg, 1997, Zelinsky et al., 1997) eye movements proved to be especially useful in reading and thus information processing in general (for a review see Rayner, 1998).

VI. Experiment

1 Introduction

1.1 Theoretical background

1.1.1 Comparative visual search

Simple tasks like making breakfast can be subdivided into smaller sub tasks, that rely either on direct visual input or on memory. Such a task in which participants have to search their surrounding in order to fulfill is called visual search task (Schmidt and Zelinsky, 2011, Hayhoe et al., 2003). The standard visual search paradigm encompasses the search for target items among distractors. Due to physical and cognitive processing limitations the instant recognition of target items can be prevented (Boot et al., 2009). By allocating the focus of attention and therefore also cognitive resources to regions of interest, such limitations can be counteracted.

There are many variations of the standard visual search paradigm. Kibbe and Kowler (2011) for example used a category search task version, in which participants had to explore arrays of hidden objects to find three multi-featured targets that belong to the same category. By varying the item grouping rule complexity and motor costs, they got new insights on cognitive and motor demands associated with both trade-off options. In contrast to normal single feature search, where items differ in one feature like color or shape, conjunctive search uses items that differ in several dimensions such as color and shape (Bacon and Egeth, 1997). These slightly more complex versions of standard visual search tasks enable insights about feature integration and chunking.

A more abstract form of this visual search paradigm is the comparative visual search task introduced by Pomplun et al. (2001). In this experimental paradigm two nearly identical display halves, also called hemifields, are presented to a participant, with only one half being visible at a time. The subjects' task is to detect the mismatch by comparing the two display halves. To assess participants' behavior eye tracking is often used. Features used in such a task are usually low-level features like color or form (Pomplun et al., 2001). The main advantage of a comparative visual search task in comparison to simpler visual search tasks is the introduction of a higher working memory involvement, as the participant has to memorize the objects within one hemifield in order to compare them to the other side (Pomplun et al., 2001).

According to Pomplun et al. (2001) not all important aspects of visual search are involved in the standard visual search task as it misses the element of memory. For the most parts, participants have to memorize one target item in order to solve the task, which is less than suited to recess the understanding of memory in visual search. By using a sequential picture matching approach however, participants have to memorize multiple objects. Usually the number of objects within a hemifield is rather large so a sequential approach is expected (Pomplun et al., 2001). They introduced this paradigm in hope that it produces new insights about the underlying cognitive processes.

1.1.2 Decision making: Behavioral trade-off

The basic idea of decision-making is, that there is more than one choice, leading to the necessity of selecting (Dolan and Dayan, 2013). A decision is usually based on several parameters. Decision making can be differentiated between reflective and reflexive. While reflective decision making is also referred to as goal-directed, model-based or prospective, reflexive decision making is often referred to as habituation, model-free or retrospective (Dolan and Dayan, 2013). Those pairs can be understood as the ends of a continuous behavioral spectrum.

While the field of decision making is ever growing Mobbs et al. (2018) stated that the next challenge is to understand the effect of trade-offs across tasks. Selecting a behavior in a certain situation one must first assess the possible behavioral options and then weight the pros and cons against each other and determine the optimal strategy to fulfill a certain goal. In such a trade-off, separately advantageous but conflicting properties or behaviors are weighted against each other to best fit a predefined goal (Del Giudice and Crespi, 2018). Goal directedness is therefore the basis for a trade-off because without a goal it would be indifferent which strategy to use. There are many disciplines that deal with cognitive trade-offs, but overarching research regarding this topic is sparse (Del Giudice and Crespi, 2018).

A certain behavior or feature is based on constraints, compromises and opposing priorities that can only be fully understood in relation to the underlying trade-off (Del Giudice and Crespi, 2018). Aspects that influence a strategy trade-off include temporal costs like delays, working memory management, energy costs like motor effort, contrast sensitivity, motor variability and probably many more (Kibbe and Kowler, 2011). One of the best-known studies regarding search strategy stability was conducted by Boot et al. (2009). In a series of experiments, they found that participants tend to have a default search strategy, which they could and would change. It is needless to say that such a strategy trade-off is highly adaptable and depends on different factors.

There are many examples for trade-offs like exploration-exploitation (Macready and Wolpert, 1998, Cohen et al., 2007, Audibert et al., 2009, Mehlhorn et al., 2015) trade-off or speed-accuracy trade-off (Wickelgren, 1977, Salthouse, 1979, MacKay, 1982). The trade-off I will go into more detail is the memorization and acquisition trade-off in the comparative visual search paradigm.

1.1.3 Memorization and acquisition trade-off in visual search

Using the comparative visual search paradigm introduced by Pomplun et al. (2001) enables the investigation of a memorization and acquisition trade-off. Using this paradigm parameters like the temporal delay and perceptual as well as cognitive complexity of objects can be alternated in order to modulate the associated trade-off and investigate determining factors.

As described before either strategy has its advantages as well as disadvantages. The management of available resources like time and working memory capacity is not a matter of favoring one strategy over the other but requires balancing (Kibbe and Kowler, 2011). Kibbe and Kowler (2011) stated that the trade-off between memorization and acquisition strategy is determined by several factors like perceptual load, cognitive demand, decision strategies and motor effort.

Especially in the comparative visual search tasks, people have to keep some internal representation in memory in order to be able to solve it (Pomplun et al., 2001). A participant has to memorize the target objects with all their features in order to compare them to the other display half. However memory is fragile so there is a larger insecurity involved by relying on it (Kibbe and Kowler, 2011). Visual details are continuously forgotten as one of the main aspects of working memory is its temporal limitation (Barrouillet et al., 2004, Barrouillet and Camos, 2012). Especially by introducing a delay the accuracy of the memory decreases, as the time period in which information has to be maintained is extended (Hardiess et al., 2011, Hardiess et al., 2008a).

However Li et al. (2018a) show the use of memory in a 3D visual search task, which suggests that the use of memory could be even greater in more natural environments. This finding is important to consider in the memorization acquisition trade-off, because it suggests that the strategy was selected to also minimize with memory associated costs. Importance of an information (Dunlosky et al., 2011), as well as emotional connotation (Kensinger and Corkin, 2003), accessibility or predictability (Menon, 1993) for example enhance memory and therefore increase the probability that the memorization strategy is used.

Another way to solve a visual search task is by the use of an acquisition strategy. In contrary to the memorization strategy, working memory load is reduced by an increased use of eye and head movements (Kibbe and Kowler, 2011). Ballard et al. (1995) showed in a Block arrangement copying task that people preferred to re-visit previously examined locations, rather than rely on memory.

In contrast to the limited working memory capacity the ability of generating eye movements is not limited directly (Ballard et al., 1995). Nevertheless, motor functions are costly regarding time and energy, so there is an indirect constraint using this strategy. These costs become spatially noticeable in the planning of saccadic eye movements, as they require time and attention and when the distances that must be travelled get larger, reducing the use of acquisition behavior when to costly (Ballard et al., 1995, Hardiess et al., 2011, Hardiess et al., 2008a, Hardiess and Mallot, 2015, Inamdar and Pomplun, 2003).

In order to perform a complex task it could be advantageous to rely more on motor behavior in order to free limited working memory capacity needed for thinking or planning and not just for memorizing available information (Kibbe and Kowler, 2011). Increasing the cognitive demands of a task as well as decreasing the motor demands resulted in more visits and a higher use of acquisition strategy contrary to a reliance on memory (Kibbe and Kowler, 2011). The effects of task complexity on search strategy could be explained by a modulation of working memory resource. Within such a paradigm the participant has to memorize visited objects, formulate new hypothesis, test them and sample new locations, which are all cognitive demanding tasks (Kibbe and Kowler, 2011). Interestingly, even though working memory capacity influences the strategy trade-off, Kong et al. (2010) found no correlation between individuals visual working memory size and solution optimality. This finding further enhances the assumption that this strategy trade-off is overlearned and optimized to fit an individual's prerequisites.

A delay, so an increase in temporal costs due to a waiting period, led to a larger reliance on memory compared to no additional temporal costs (Kibbe and Kowler, 2011). With a temporal delay, participants tend to search slower and more thorough in order to avoid the delay. An explanation for this finding proposed by (Kibbe and Kowler, 2011) is that the time it takes to perform an action is viewed as a resource that needs to be managed. The time that was spent waiting for the delay to pass is unproductive and was therefore avoided. Alternatively, this could reflect the knowledge of memory loss with a longer delay, which was counteracted by searching more thorough.

1.2 Synthesis of current state of the art and present study

In order to extend the body of evidence regarding working memory usage and strategy selection the present study was conducted. The aim of this study will be to investigate strategies associated with chunking. To do so a comparative visual search task adopted from Pomplun et al. (2001) was used.

In addition, different delay and complexity conditions were included in order to modulate the associated costs. Furthermore, a second experimental Block was included in which participants had to further change their search strategy, due so a limitation in the number of hemifield switches available. This behavioral investigation based on perceptual and cognitive costs will also be discussed.

Behavioral and eye tracking data of this experiment was collected within my bachelor thesis (Bräutigam, 2018), but the eye tracking data was not investigated previously. By examining eye movements in detail, we hope to gain some new insights not obtainable with behavioral data only.

The present study will investigate macro levels of attention allocation, in which successive fixations are representative of attention shifts (Scinto et al., 1986). In this experimental design micro attention shifts were not controlled and therefore disregarded. Even though attention shifts can be covert or overt, only overt attention will be examined as there is a direct interaction between eye movement and focus of attention (Findlay et al., 2003).

1.3 Hypotheses

First of all, we hypothesize that the eye tracking data will be comparable to the behavioral data. To assess this hypothesis three parameters will be compared: Trial duration, number of switches and lastly the time per switch.

Secondly the classification method used to distinguish between object and non-object fixations will be tested. We assume that the used methods will lead to a conservative object classification but will have an overall good performance.

The following hypotheses will consider the eye tracking data itself. We assume that there will be a strategy change between the first and the second Block, visible mainly in the number of switches, total fixation time and number of switches, similar to the findings acquired from behavioral data.

Due to the task requirements, there will be fewer switches in Block 2, which participants had to counteract by changing their acquisition behavior. This adjustment will be particularly noticeable in harder delay and feature combinations.

Furthermore, we predict that trials with complex objects will be harder for participants to solve, which will be visible in the fixation duration of such objects, the number of fixations, the total fixation time and the number of hemifield switches. Similar effects, however less pronounced will be visible for the long delay condition.

Additionally, we predict that there will be a difference between the two screen sides due to the difference in cognitive processes involved. This difference will probably be most visible in the total fixation time. Participants will most likely use the left side as a template to compare the right side to.

Lastly, we anticipate that the classification of fixations as chunk or revisit will lead to new insights regarding cognitive processes within a comparative visual search task. We hope that including this differentiation will explain why some findings previously expected could not be found by just looking at the behavioral or eye tracking data without this distinction. Furthermore, chunk size should resemble working memory capacity.

2 Material and Methods

2.1 Participants

In total there were 18 Participants age 19 to 23 (21.06 ± 1.18 years). However, two participants had to be excluded prior to the analyses due to poor execution and high error rates. After the data overview another participant was excluded due to non-usable eye tracking data in Block 2. All participants were naïve to the experiment and had normal or corrected vision. A written informed consent was obtained from each participant.

2.2 Setup

In a controlled tempered and air-conditioned room, participants sat 50 cm in front of a PC. The used computer screen was a standard 1280 x 1024 px, 60 Hz monitor. The computer was running Windows 2010. The experiments were presented, controlled and recorded using MATLAB with the Psychophysics Toolbox extensions from Brainard (1997).

Eye movements were recoded using a monocular 'remote' system with a sampling frequency of 60 Hz (Eyegaze System, LC Technologies, Inc.). The participants head was fixated using a head and chin rest.

2.3 Procedure

Participants were placed in the experimental setup, the infrared contrast between iris and pupil was determined and the eye position was controlled. For the comparative visual search experiment the eye tracking device had to be calibrated first, which was done by following randomly appearing blue circles on the screen.

After explaining the tasks participants started with a test trial, after which the experiments began. Participants task was to compare the two screen sides as quickly and reliable as possible and then report the number of differences between the two sides verbally. All trials from a Block were in succession but participants could incorporate breaks if needed. In order to control for unwanted rehearsal processes and restrict memory processing to the visual representations, participants had to repeat three syllables 'bla-bli-blub'.

Each trial started with the written presentation of the condition combination. After a 2.5 s delay the objects on the left screen side were visible. The right side was masked. With a mouse klick the mask could be switched from left to right or the other way round. With a klick both screen sides were masked for the duration of the delay (Figure 1). A trial could be terminated by clicking the space bar. At the end of a trial participants had to report the number of differences, before starting a new trial.

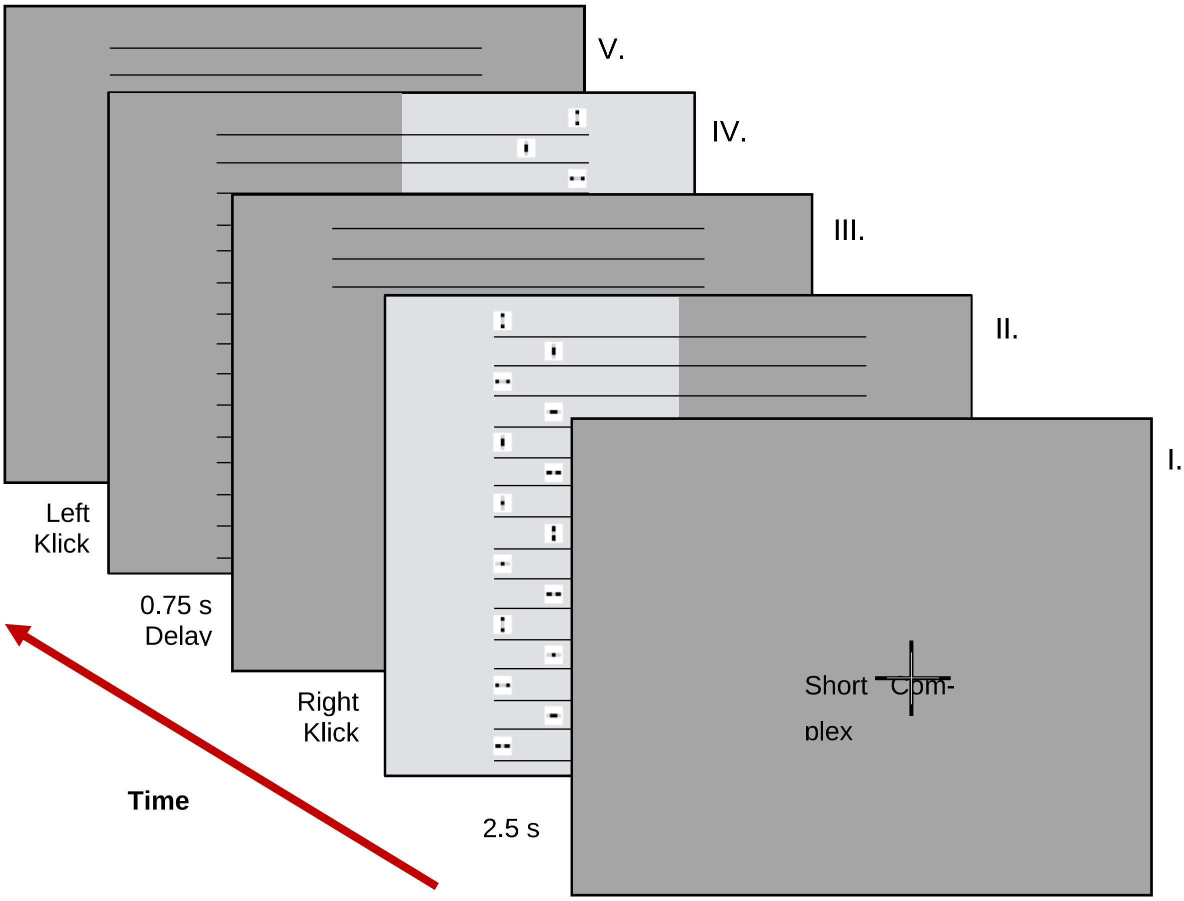

Figure 1: Trial procedure

I. *Presentation of the fixation cross(2.5 s) and trial condition: ‚short delay' (0.75 s) and 'complex features' (three features). The text was masked after 1s, making only the fixation cross visible for another 1.5 s.*
II. *Stimulus presentation with initial masking of the right side and isochronous guiding lines.*
III. *After right klick: the whole screen is masked for the delay duration. In this case 0.75 s. Guidelines stay visible*
IV. *Stimulus presentation right side. Left side is masked. Isochronous guiding lines are visible.*
V. *After left klick: the whole screen is masked for the delay duration. In this case 0.75 s. Guidelines stay visible.*

In Block 1 the number of switches depending on the condition was recorded and used in Block 2 to restrict acquisition behavior. In Block 2 there was a reduction of the number of usable screen switches to 75% of used switches in Block 1. This reduction depended on the four used conditions. The number of switches still usable was displayed in the middle of the screen in red in the area of the fixation cross.

Both Blocks consisted of 20 trials (2 delay conditions x 2 complexities x 5 repetitions), making a total of 40 trials. Within those trials eye tracking and behavioral data were recorded.

2.4 Conditions

In the main experiment two variables were varied: first the delay between the screen side switches and second the object complexity. With two parameter values in each case this resulted in a two-by-two condition matrix (Table 1).

Table 1: Condition matrix with corresponding colors
Two delay durations: short (pink) and long (purple) and two object complexities: simple (blue) and complex (yellow)

		Delay	
		Short (0.75 s)	Long (1,5 s)
Complexity	Simple (2 features)	‚Short Simple‘	‚Long Simple‘
	Complex (3 features)	‚Short Complex ‘	‚Long Complex ‘

The delay between switches was either short with 0.75 s later visualized in blue or long with 1.5 s later visualizes in yellow.

The object complexity was either easy (pink) consisting of just two features color and orientation or complex (purple) consisting of three features. The three object features used in the complex condition were color, orientation as well as gap size.

2.5 Objects

In this experiment 12 different objects were used in order to suffice the two objects complexities. Each object was embedded in a white 40 x 40 px square. The four objects used in the simple condition consisted of an either black or grey bar in vertical or horizontal orientation. The eight other objects were made up of a combination of 1) orientation: vertical or horizontal, 2) color of the end pieces: black or grey and 3) gap size small (8 px) or large (16 px) (Figure 2). An in-detail depiction of a complex object with all dimensions can be found in the supplementary: Figure 68.

In each hemifield 15 objects were placed in two rows. An equal spacing of objects reduced the effects of spatial proximity as a grouping factor and enabled us to better assign a fixation to an object.

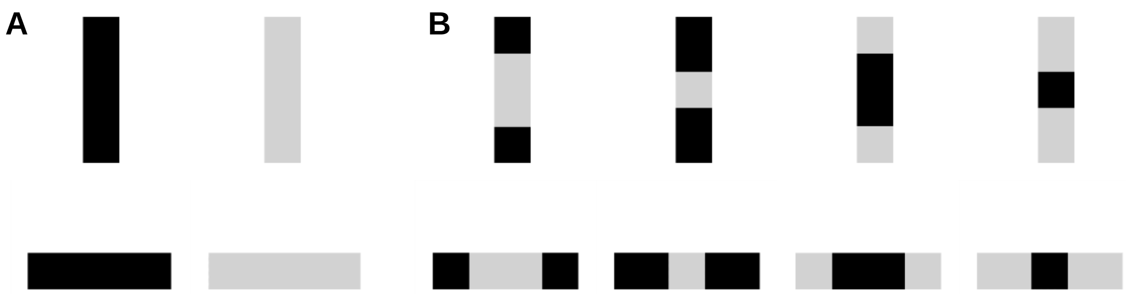

Figure 2: All 12 used objects (Original: 40 x 40px)
A: Pictures of all four feature combinations with two features: 1) vertical and horizontal orientation and 2) colors black and grey.
B: Pictures of all eight feature combinations with three features: 1) vertical and horizontal orientation 2) colors of the ends black and 3) grey and gap size small and large.

2.6 Data structure

Resulting from the eye movement recordings the data consisted of x and y coordinates for each time point within the sampling period. From this dataset fixations were extracted using fixed definitions. A fixation was defined by the eye movement speed as well as the duration. Eye movement speed had to be less than 16.666 °/sec. The fixation duration was defined as at least 120 milliseconds. Usually fixation durations vary vastly between hundred milliseconds to more than a second (Trukenbrod and Engbert, 2014).

Eye movement trajectories were discarded. As a result, the only the duration of each fixation as well as the fixation position with x and y coordinate were used. From this data structure four parameters could be derived: the fixation duration, the number of fixations, the overall Fixation duration and fixation position. From the fixation position switches can be inferred. However, by just having the fixation coordinates examinable hypotheses are limited. Therefore, fixations were classified as an object fixation with the corresponding object number or non-object fixation. The method of this classification will be discussed and evaluated in detail in chapter 2.9: Evaluation of eye tracking data and classification method.

All these resulting parameters can be examined either on trial level, similar to behavioral data, or on switch level. As we will see later, looking at both levels enables a deeper understanding of the underlying mechanism. While trial level gives an overview, switch level resembles a cognitive processing step and as Pomplun et al. (2001) put it are of particular relevance for a detailed investigation of the strategies in a comparative search tasks. The biggest advantage of the possibility to assess a processing step is the in-detail examination of fixations.

Following Pomplun et al. (2001) successive fixations within a hemifield will be examined in detail. Due to the use of eye tracking and the deep going investigation of switch level, we can classify fixations as either *chunk*: the number of different fixated objects within a switch or as a *revisit*: a refixation of an already visited object within one switch. By determining the number of fixations classified as chunk we can approximate the chunk size, which gives an approximation of working memory capacity as it resembles the number of different objects a participant could hold in memory within a processing period. The number of revisits on the other hand can be seen as a way to consolidate the acquired information before switching sides and comparing them to the corresponding objects.

Figure 3 illustrates why fixations were analyzed on the basis of a switch later on and how the distinction between chunk and revisit fixation was made. In this example participant 1 started at the fixation cross and then fixated the first object. After fixating the second and third object all three previously fixated objects were revisited. In this example the chunk size would be three and the number of revisits would be three as well. After six object fixations he switched sides and fixated objects on the right.

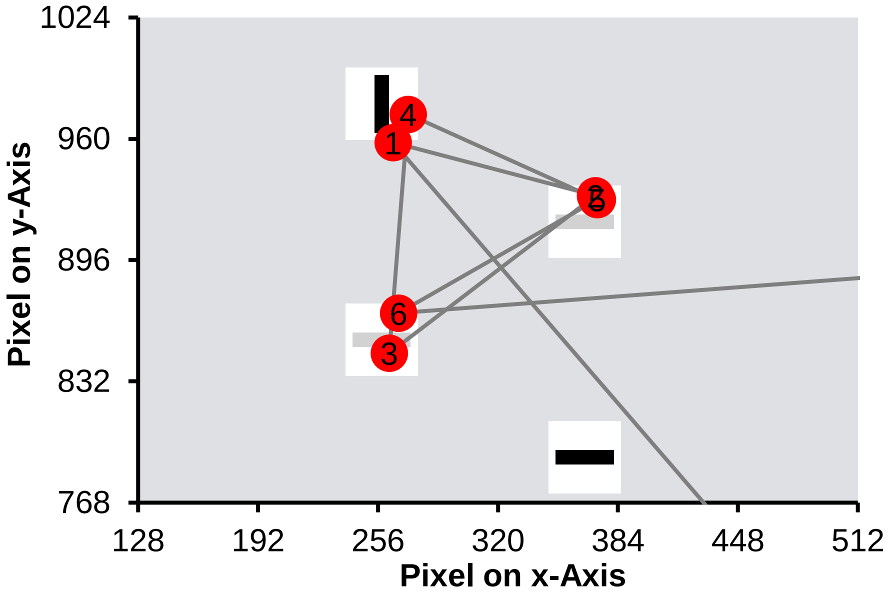

Figure 3: Example of all object fixations within the first switch
Example from participant 1 trial 8 Block 1. The fixations are plotted using x and y pixel coordinates. A fixation is visualized as a red dot and numbered, while the eye trajectory is plotted as a grey line.

2.7 Statistical analysis

The experiment was constructed in a within participant design, making adequate statistical analysis necessary. Repeated measurement ANOVAs as well as paired t-tests were the chosen analytical methods.

If not specified otherwise the data was normally distributed. Normal-distribution was analyzed using Kolmogorov-Smirnov-tests in IBM SPSS statistics. Even though it would be better to use non-parametric tests if the distribution was significantly different from normal, parametric tests such as paired t-tests for the difference between the Blocks, Pearson correlations as well as three-factorial repeated measurement ANOVAs for the effects of different variables were used. The reason for this decision is the importance of individual differences and the large number of repetitions for each participant. The large individual differences do not result from errors but rather reflect the individual strategies and should therefore be considered.

Due to the experimental design, there were only two manifestations within the analyzed variables, making it impossible and unnecessary to test for sphericity.

In figures significance levels are visualized in the following way:

$p \leq 0{,}05$ = *

$p \leq 0{,}01$ = **

$p \leq 0{,}001$ = ***

For participants 9 and 15 the eye tracking data could not be used due to technical difficulties, reducing the number of participants to 16. Furthermore, the pupil could not be detected in Block 2 for participant 12, making his exclusion necessary as well.

Most fits have intercept at the origin due to logical deduction. For most fits a zero in one variable like fixation duration would first of all not be possible due to task constraints and second correlated number of fixations should also be 0 as a consequence.

Statistical outliers are defined as 1.5 times the interquartile range (IQR).

2.8 Comparison behavioral and eye tracking data

In order to compare whether eye tracking data is similar to behavioral data three variables were analyzed: trial duration, number of switches and time per switch.

2.8.1 Trial duration

The first parameter we can compare is the trial duration. It is important to state, that the trial duration for eye tracking data was reconstructed using the duration of object fixations. On the other hand, the trial duration of the behavioral data was reconstructed by using the time between switches as well as the delay condition for each trial.

With an average of 44.96 ± 3.18 s for a trial for the eye tracking data and 45.94 ± 3.56 s for a trial for the behavioral data there was no difference between the two methods (two-sided paired t-test: $t(14) = -1.59$, $p = 0.13$;Figure 4: A). The trial duration was highly correlated between the behavioral and the eye tracking data (Pearson: $r(15) = 0.99$, $p < 0.001$, Linear fit: $y = 1.03x$, $R^2 = 0.998$; Figure 4: B).

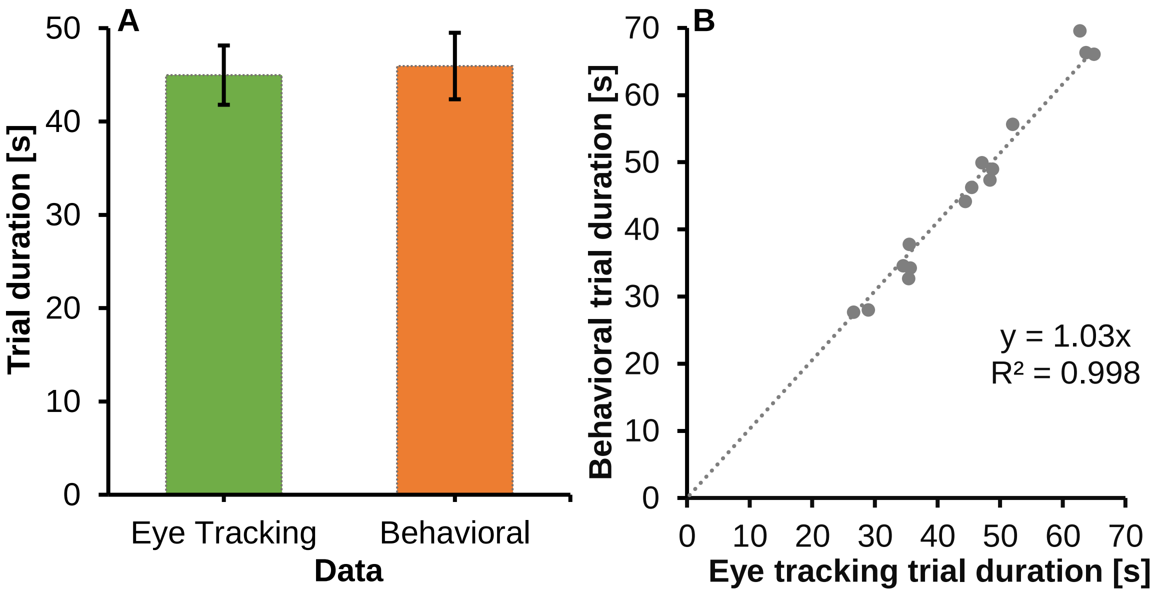

Figure 4: Comparison of the trial duration between eye tracking and behavioral data
A: Trial duration [s] is plotted against the data type: eye tracking (green) and behavioral (orange) averaged over participants (n = 15).
B: Correlation between the trial duration [s] averaged for a participant using either eye tracking (x axis) or behavioral data (y axis).

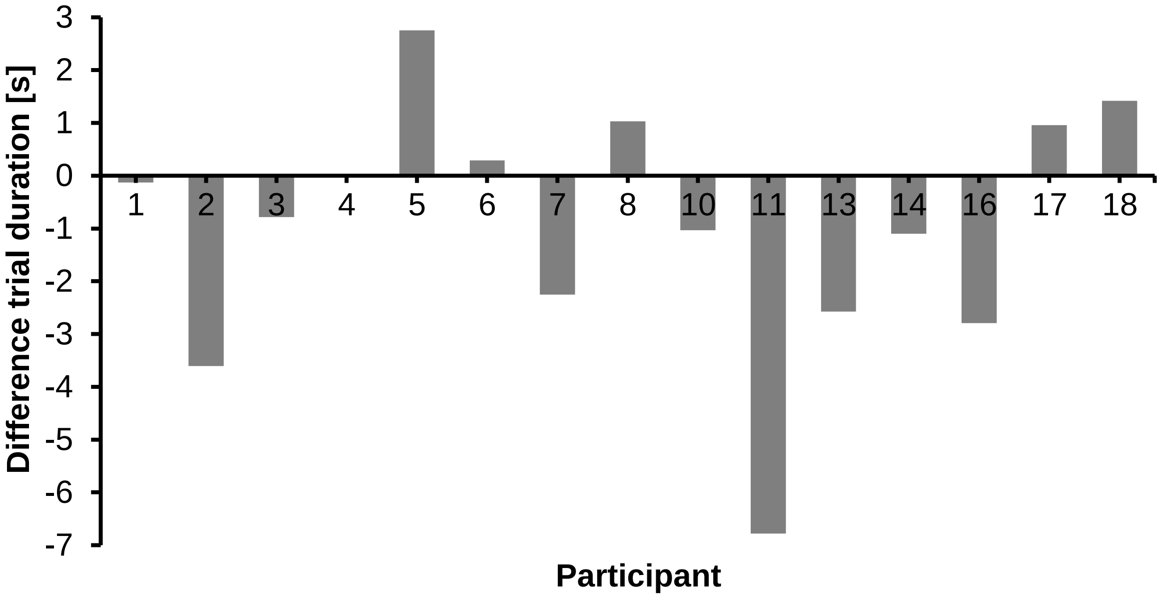

Figure 5: Difference of trial duration between behavioral and eye tracking data for participants
Differences of trial duration [s] between eye tracking and behavioral data plotted against the participant. Negative differences represent that a trial was longer using behavioral data, positive values implies that a trial was longer using eye tracking data.

By looking at the difference in trial duration in more detail we can see that participants differed vastly. For example, a trial for participant 11 took on average 6.78 seconds more when looking at the behavioral data than the trial duration reconstructed by fixation durations suggested (Figure 5). On the contrary the mean trial of participant 5 took 2.75 seconds longer using eye tracking data than a trial was suggested by the behavioral data.

As expected, the mean trial duration was slightly longer for the behavioral data. However, this difference was rather small, suggesting that the participants fixated objects nearly the whole time a trial was running. The detailed analysis showed that there were some unexpected discrepancies: The trial duration reconstructed from the fixation durations should not be longer than the behavioral data suggests. A reason for this anomaly could be that the start eye tracking in the beginning of a trial varied slightly. While the behavioral data started the time measurement as soon as the objects where visible, the start of a trial in the eye tracking data was defined by the first fixation to an object position. However, it cannot be excluded that participants started fixating an expected object position even though there was no object visible because they knew were an object will become visible.

2.8.2 Number of Switches

In order to compare the number of switches a switch had to be defined. For the behavioral data a switch was a klick, which triggered a mask to switch from one side to the other. A switch in eye tracking data was defined by a fixation to an object on the other side of the screen.

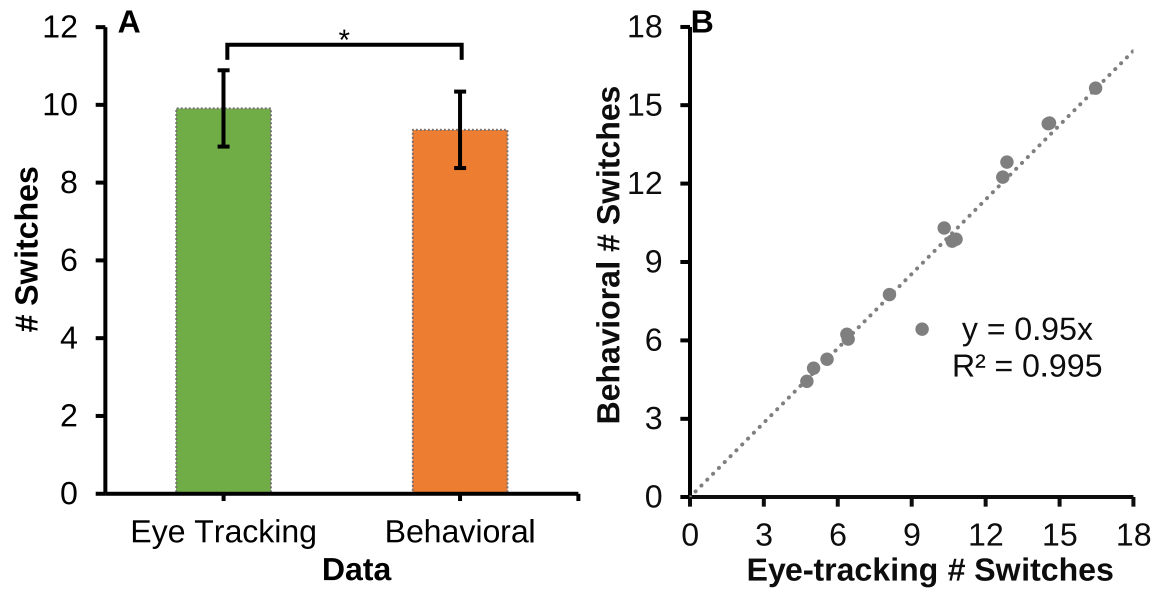

Figure 6: Comparison of the number of switches between eye tracking and behavioral data
A: Number of switches is plotted against the data type: eye tracking (green) and behavioral (orange) averaged over all participants (n = 15).
B: Correlation between the number of switches using eye tracking (x axis) and behavioral (y axis) data averaged for each participant (n = 15).

There were a total of 5946 switches using eye movements and 5616 switches using klicks. This corresponds to an average of 9.91 ± 0.98 switches per trial using eye movements and 9.36 ± 0.98 switches per trial using klicks (Figure 6: A). Comparing this on an individual level yields a difference between eye tracking and behavioral data regarding the number of switches (Paired two-sided t-test: $t(14) = 2.99$, $p = 0.12$). However, by looking at the correlation between those two measurements, we can see that with an increasing number of switches in the eye tracking data the number of switches in the behavioral data also increased (Pearson correlation $r(15) = 0.98$, $p < 0.001$, linear fit: $y = 0.95x$, $R^2 = 0.995$; Figure 6: B).

The mean differences in switches between eye tracking and behavioral data ranged from an average of 0.025 switches more in the eye tracking data for participant 8 to an average of 3 more switches using eye tracking data compared to behavioral data for participant 13 (Figure 7). Since participant 16 had the biggest difference between eye tracking and behavioral data regarding the number of switches excluding this participant lead to a nearly perfect correlation (Pearson correlation $r(15) = 0.997$, $p < 0.001$, linear fit: $y = 0.96x$, $R^2 = 0.9993$).

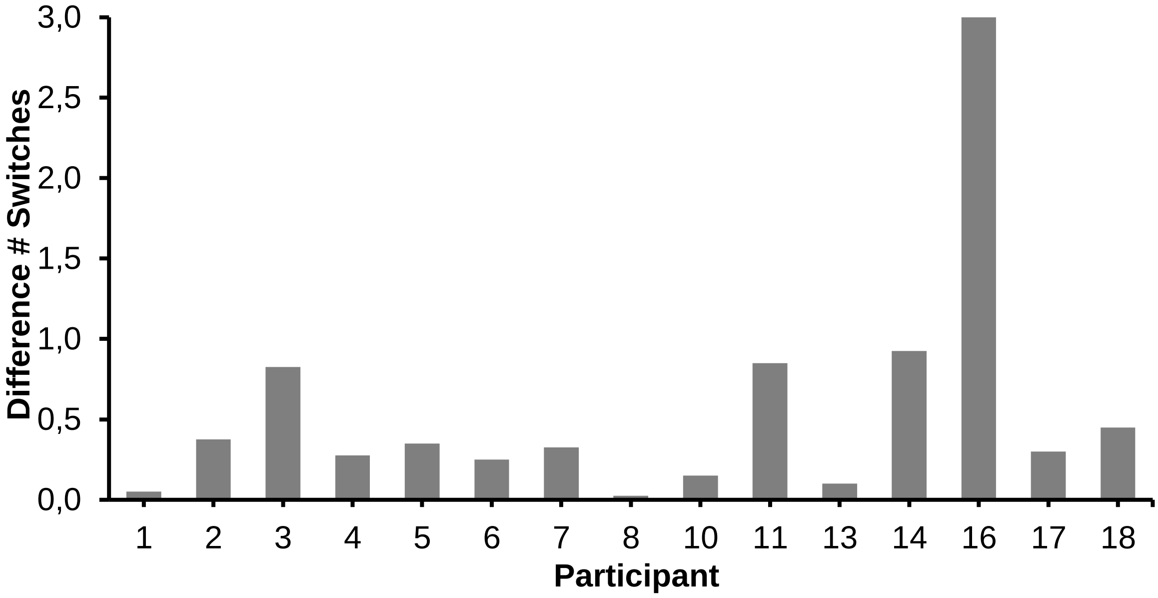

Figure 7: Differences of number of switches between eye tracking and behavioral data for participants
Differences of number of switches between eye tracking and behavioral data plotted against the participant. Negative differences represent more klicks than hemifield switches using the eye movements, positive values mean more hemifield switches using eye movements than klicks.

Comparing the number of hemifield switches using the number of klicks as a behavioral measurement and the eye movements as an alternative measurement it can be seen, that in 445 of a total of 600 trials (pooled over all participants) both measures coincided. The differences in the number of switches ranged from 2 more klicks than eye movement switches to 16 more eye movement switches than klicks. In general, only 2 trials had more switches in the behavioral data than in the eye movement data and 153 trials had more eye movement switches than klicks (Table 2 and Figure 8). However, from those 153 trials with more eye movement switches 82 had one more eye movement switch and 40 had two more eye movement switches.

Table 2: Overview of differences between behavioral and eye tracking data
Difference between eye tracking and behavioral data regarding the number of switches and the number of observations for each difference. Negative differences mean more klicks than hemifield switches using the eye movement, positive values represent more hemifield switches than klicks (n = 600 observations).

Difference	**-2**	**-1**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**8**	**10**	**16**
# Oberservations	1	1	445	82	40	10	8	1	7	1	1	2

Even though there was a statistical difference in the number of switches between eye tracking and behavioral data most of the trials had the same or nearly the same number of switches. Participants tended to make one more switch with the eyes compared to klicks. The nearly perfect correlation as well as the linear fit however, enhances the hypothesis that eye tracking and behavioral data are both adequately accurate. Participant 16 had on average 3 more eye movement switches compared to klicks, which could be an indicator for spatial rehearsal.

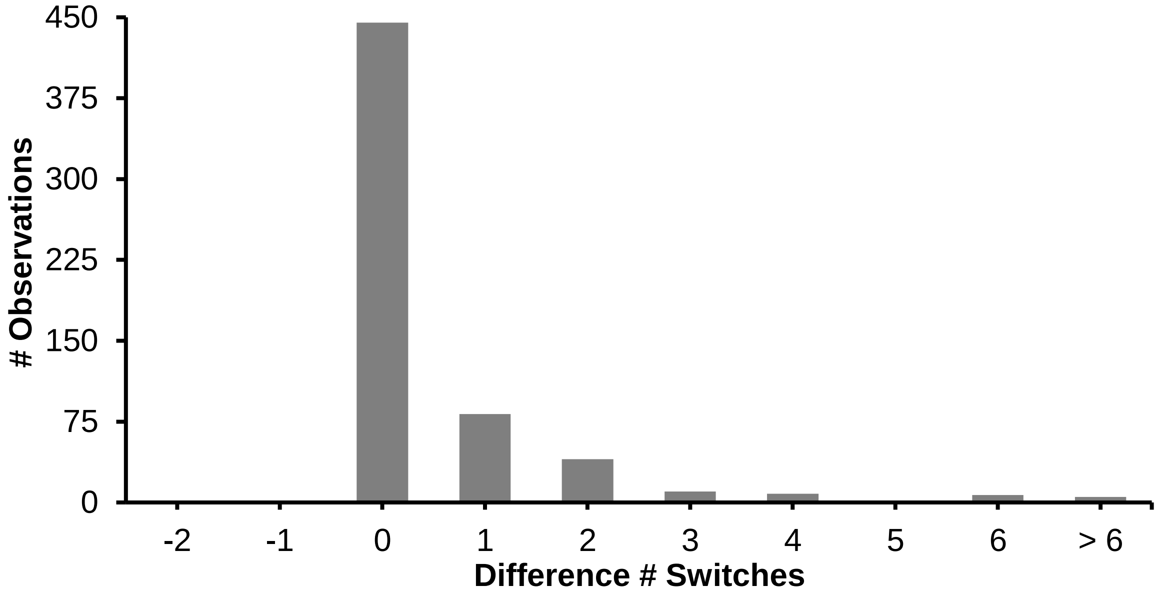

Figure 8: Difference of number of switches between behavioral and eye tracking data
Number of observations plotted against the difference in the number of switches between behavioral and eye tracking data. Negative differences in the number of switches mean more klicks than hemifield switches using the eye movement, positive values represent more hemifield switches using eye movements than klicks. Data was pooled over a total of n = 600 observations.

2.8.3 Time per switch

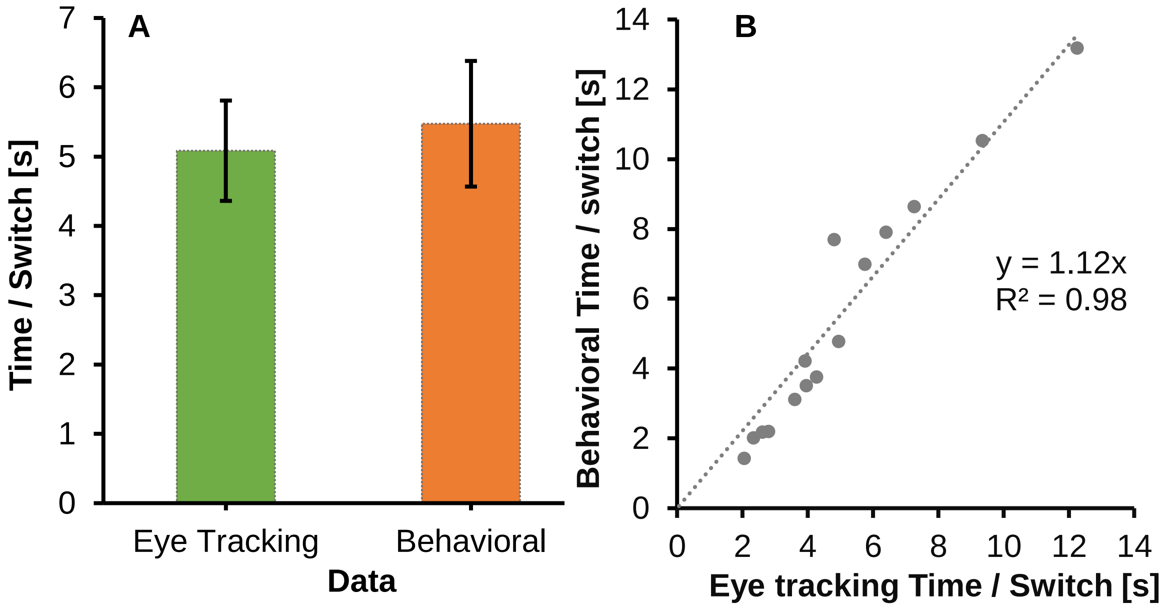

Figure 9: Comparison time per switch for behavioral and eye tracking data
A: Boxplot of time per switch [s] plotted for the data: eye tracking (green) and behavioral (orange) (n = 15)
B: Correlation between the time per switch [s] averaged for the participants using eye tracking (x axis) and behavioral (y axis) data. The last parameter that can be compared between eye tracking and behavioral data is the time per switch. The same definition for a switch as before was used.

The mean of total fixation time per switch was 5.08 ± 0.72 s and ranged from 2.05 s for participant 4 to 12.25 s for participant 12 (Figure 9: A). The mean duration of a switch in the behavioral data was 5.27 s and the durations per switch ranged from 3.11 s in participant 1 to 13.18 s in participant 13. There were no significant differences between eye tracking and behavioral data (Paired two-sided t-test: $t(14) = -1.42$, $p = 0.18$). The time per trial increased continuously for eye tracking and behavioral data (Pearson: $r = 0.968$, $p < 0.001$, $n = 15$, linear fit: $y = 1.12x$, $R^2 = 0.98$; Figure 9: B).

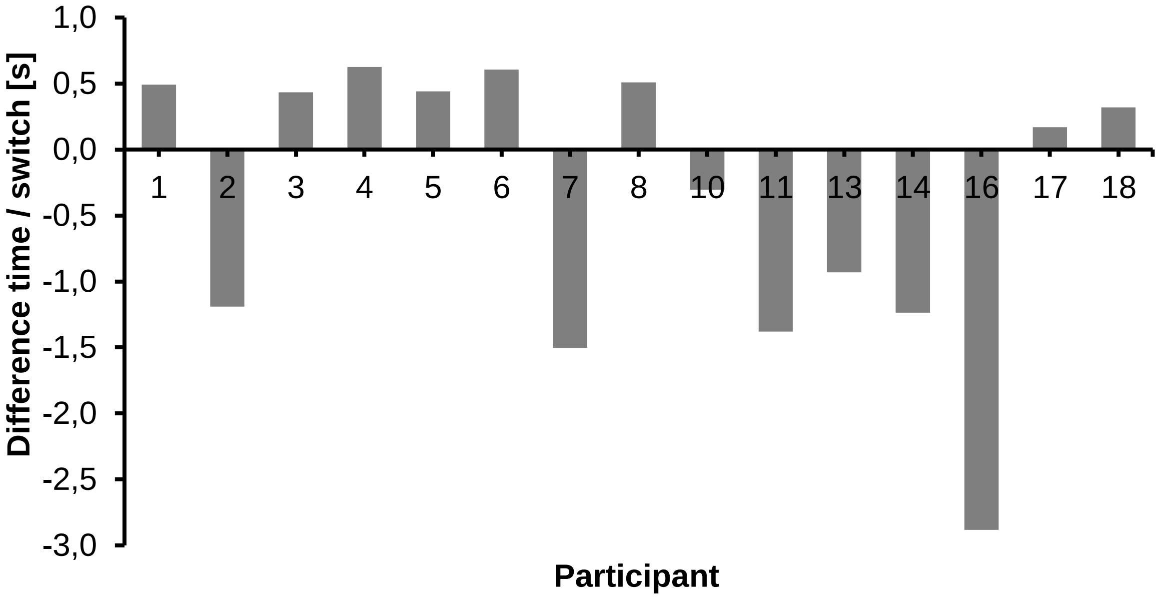

Figure 10: Differences of time per switch between eye tracking and behavioral data for participants Differences of time per switch [s] between eye tracking and behavioral data plotted against the participant. Negative differences represent longer times per switch of behavioral data, positive values mean longer times per switch for eye tracking data.

Going more into detail, we can see that the time per switch was longer in the behavioral data for most participants compared to eye tracking data. This mean difference was the biggest with 2.88 s longer in the behavioral data for participant 16. Participant 4 on the other hand took approximately 0.44 s longer per switch, when eye tracking data was used (Figure 10). However, it is to note that most of the differences are not longer than 1.5 s.

An explanation for those findings is that participants did not only fixate objects within a switch, they also did saccades and non-object fixations, which were not accounted for in the eye tracking data. Another explanation would be that in the behavioral data the delay period was not included in the time per switch because there was nothing to look at.

Despite all those problems the median of time per switch was nearly the same for both methods reinforcing the hypothesis that eye tracking and behavioral data could both be used in such a paradigm.

2.9 Evaluation of eye tracking data and classification method

In order to assign the fixations to an object half the distance between the objects was used as the radius for a catching area around an object. By using this method, it was expected that most fixations could be assigned properly and that the fixations that could not be assigned to an object were mainly due to the ambiguousness of the fixation.

Using this method from 62454 fixations recorded over all 15 participants and trials, 54098 fixations could be assigned to the 30 objects, while 8356 fixations could not be assigned to an object. With only 13.38% of all fixations not assigned to one of the 30 objects, the method of matching fixations to objects seems promising (Figure 11).

Going more into detail with the non-assigned fixations, these fixations were assigned into different areas. A visualization of those areas and an example for participant 1 Block 1 trial 1 can be found in Figure 12 and the corresponding overview can be found in Table 3.

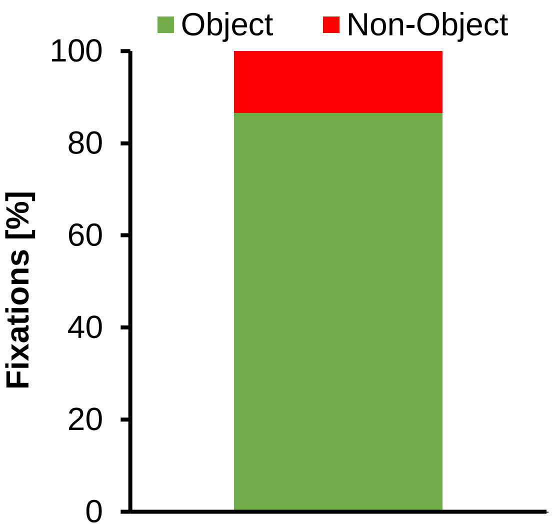

Figure 11: Fixation Overview
The proportion of fixations [%] assigned to an object (green) and to no object (red) in relation to the total number of fixations (n = 64911) pooled over all participants (n = 15) and trials (n = 40).

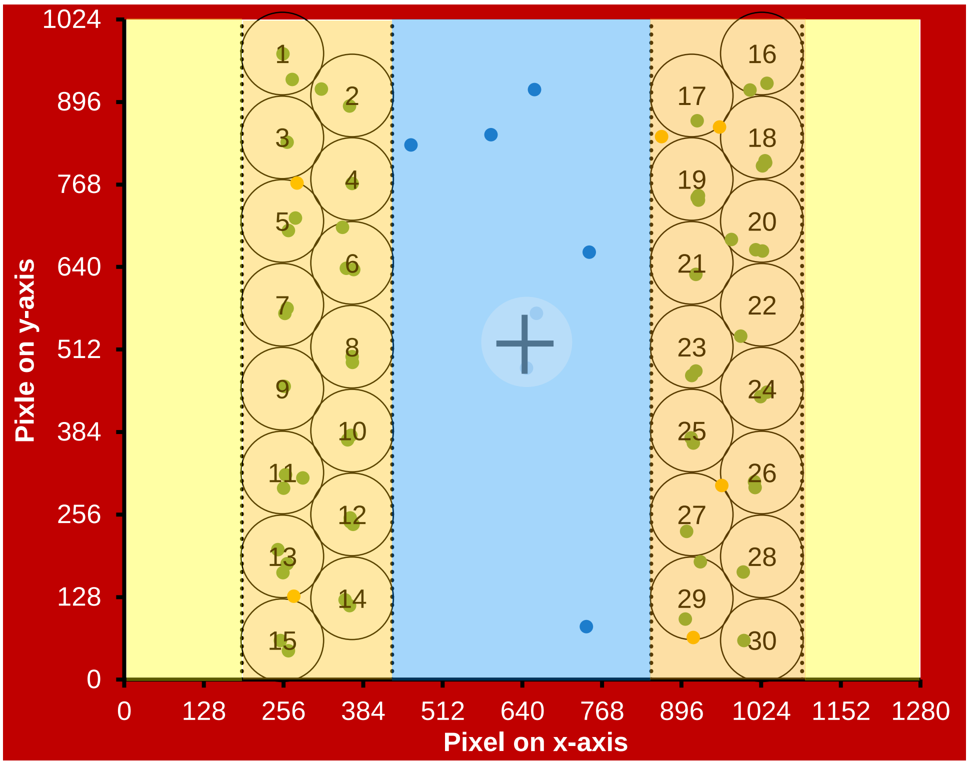

Figure 12: Example of screen areas with fixation data: Participant 1 Block 1 trial 1
Colors resemble the different areas: out of screen (red), borders (yellow), catching area left (light yellow), catching area right (dark yellow), fixations within the middle area (blue) divided into fixations within the fixation-cross area (light blue) and fixations in the middle outside of the fixation-cross area (dark blue).

Out of screen fixations (dark red), as the name suggests are fixations that were not directed to the 1280 x 1024 pixel screen. Fixations classified as on the borders (yellow) ranged from the left or right end of the screen to the corresponding catching area. The catching areas (yellow) ranged from the furthest left point within the object areas to the furthest right point within the object areas but were not assigned to one object due to a lack of unambiguousness. Most of the non-Object fixations were directed towards the middle area (blue), that's why fixations were subdivided into fixations within the fixation-cross area (light blue) and fixations in the middle outside of the fixation-cross area (dark blue).

Table 3: Overview over the total number of fixations Number of fixations in a certain area pooled over all participants (n = 15) and trials (n = 40).

Fixation area	**Number of Fixations**
All fixations	62454
Object classified fixations	54098
Non-object classified fixations	8356
Out of screen	369
Borders	143
Left Border	91
Right border	52
Middle area	5595
Fixation-cross	2294
Catching areas	2248
Left catching area	1225
Right catching area	1023

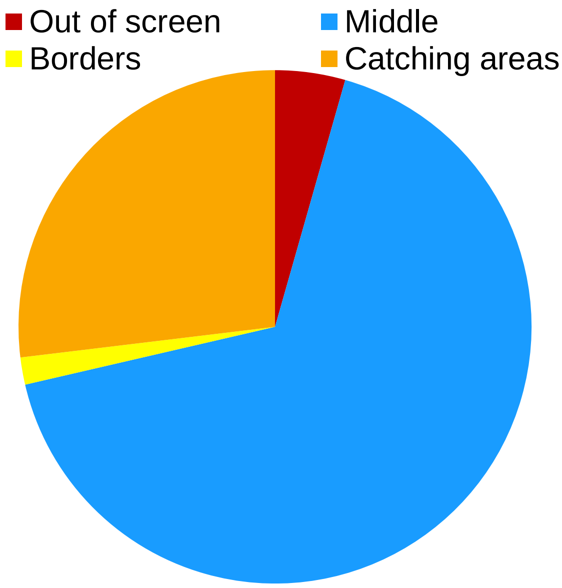

Figure 13: All non-object fixations visualized. Colors resemble the different areas: out of screen (dark red), borders (yellow), catching area (yellow) and fixations within the middle area (blue).

With those areas defined it could be seen that from the non-assigned-fixations only 368 fixations (4.42 %) were outside of the screen (Figure 13). This suggests a good involvement of participants in the tasks due to only a small portion of eye movements away from the screen.

26.90% of non-object fixations where directed to the catching areas. This is the result of the method used to classify a fixation. In Figure 12 however one can see that the fixations within the catching areas are hard to classify unambiguously, making a classification hard. Because we used catching areas around the objects with half the distance between two objects as the radius there are some fixations which are not assigned to an object even though it was in proximity of one or more objects.

Whilst only 1.71 % of non-object fixations were directed to the borders, 66.96 % of the non-object fixations where directed to the middle of the screen. A possible explanation could be that participants made a fixation during a hemifield saccade.

Fixations to the middle and the catching areas were investigated further because they made up most of the non-object fixations. Regarding the 5595 fixations to the middle, 41

% of those fixations went to the area of the fixation-cross, while the other 59 % went to positions in the middle area other than the fixation-cross (Figure 14: A). Two possible explanations for this observation are that participants fixated longer on the fixation cross, even though the trial had already started or that they checked the middle while switching screen sides. 54.4 % of all non-object fixations to catching areas went to the left, so there was no significant side bias (Figure 14: B).

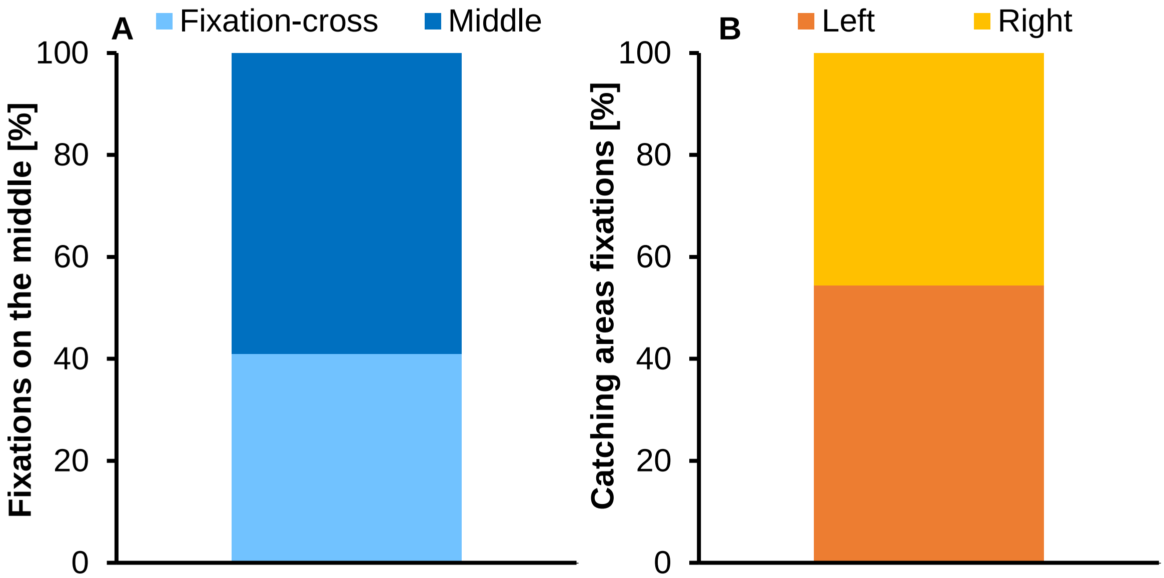

Figure 14: Fixations in detail pooled over all participants (n = 15) and trials (n = 40)
A: Relation of fixations on the middle [%] subdivided into fixations to the fixation-cross (light blue) and areas in the middle (dark blue).
B: Relation of fixations on the catching areas [%] subdivided into fixations to the left (yellow) and right (orange).

This classifying method was rather conservative but used, in order to make sure each fixation was unambiguously directed to one object. In a controlled search task by Schneider and Shiffrin (1977) fixation points were allocated selectively, mostly on one object at a time. They suggested that in such a controlled search attention is required, which would also apply to the present study.

3 Results

To examine what eye tracking data can contribute to the question of working memory capacity as well as strategy trade-offs in a comparative visual search task, we can use three parameters: the mean fixation duration, the number of fixations as well as the total fixation time, which results from the fixation duration and the number of fixations.

In addition to trial level those parameters can also be investigated within a switch. The level of a switch resembles a cognitive cycle, consisting of encoding and memorization of objects.

Because eye tracking was used, we can classify fixations by type either as *chunk*: the number of different fixated object within a switch or as *revisit*: a refixation of an already visited object within one switch. By determining the number of fixations classified as chunk we can approximate the chunk size, which gives an approximation of working memory capacity as it resembles the number of different objects a participant could hold in memory within a processing period.

First, we will investigate general measurements like the fixation duration, the number of fixations per trial as well as the total fixation time. After this we will examine the number of switches as well as the fixation behavior within a processing cycle. In this step will look at the differentiation between chunk and revisit fixations.

3.1 Fixation duration

Starting with the fixation duration we can see that there was no significant difference between the mean duration of an object fixation (0.48 ± 0.041 s) and a non-object fixation (0.44 ± 0.027 s) (Paired two-sided t-test: t (14) = 1.407, p = 0.19). The participants mean duration of an object fixation ranged from an average of 0.30 ± 0.007s for participant 10 to an average of 0.91 ± 0.06 s for participant 8 (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean duration of an object fixation: F(14) = 150.61, p < 0.001, eta-square = 0.915). In contrast the mean duration of a non-object fixation ranged from an average of 0.30 ± 0.009 s for participant 11 to an average of 0.65 ± 0.037 s for participant 10 Figure 15. Additional information can be found in Chapter VIII. Supplementary: Figure 69.

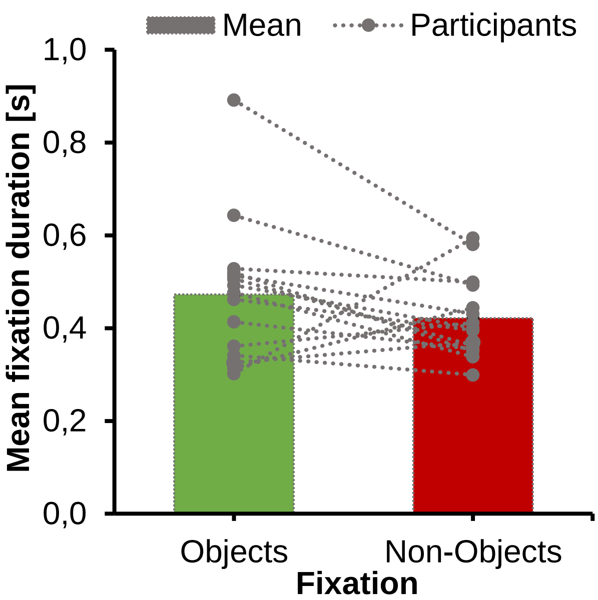

Figure 15: Mean fixation duration of object and non-object fixations
Mean fixation duration [s] plotted against the fixation: Object (green) and Non-Objects (red), averaged over all participants (n = 15) as bars as well as for each participant averaged over all fixations as dots.

In order to assess the effects of Block, complexity and delay on the mean duration of an object fixation, further analysis was conducted. It is to note that the mean duration of an object fixation from Block 2 within the complex condition was not distributed normally for either the short nor the long delay (Kolmogorov-Smirnov: D(15) = 0.295, p = 0.001; D(15) = 0.293, p = 0.001).

The duration of an object fixation differed between the two Blocks (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean duration of an object fixation: F(14) = 12.033, p = 0.004, eta-square = 0.46).

In Block 1 the mean duration of object fixations was 0.52 ± 0.044 s per fixation, while in Block 2 it was 0.44 ± 0.036 s per fixation (Figure 16: A). Furthermore, there was a correlation between the Blocks regarding the duration of an object fixation (Pearson correlation two-sided: $r = 0.85$, $p < 0.001$, $n = 15$, linear fit: $y = 0.83x$, $R^2 = 0.97$; Figure 16: B).

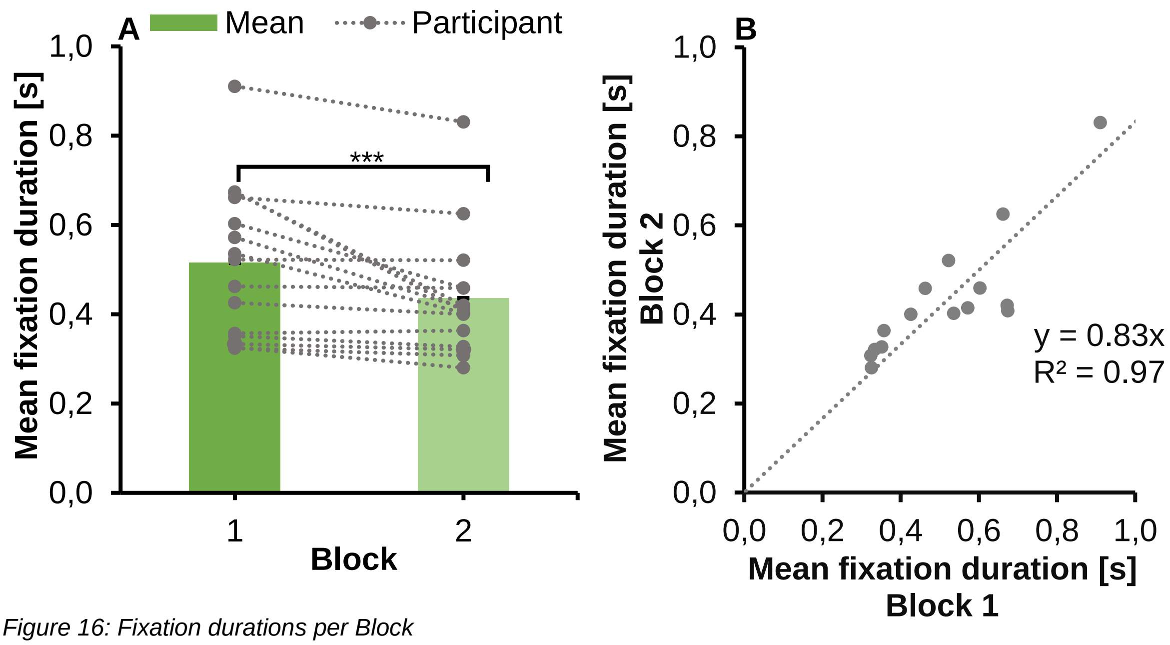

Figure 16: Fixation durations per Block
A: Mean fixation duration [s] averaged over participants (n=15) plotted against the Block for objects: green and non-objects: red.
B: Mean fixation durations [s] in Block 2 are plotted against the duration of object fixations [s] in Block 1 for n = 15 participants. A linear fit going through the null point was computed and plotted.

Even though the difference is not statistically significant a tendency could be seen: the mean duration of fixation of a complex object was with 0.51 ± 0.052 s slightly longer than the mean fixation duration on an easy object, which was 0.44 ± 0.027 s (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean duration of an object fixation: $F(1, 14) = 4.48$ $p = 0.053$, partial eta-square 0.242; Figure 17: A). For every second participants fixated an easy object they fixated a complex object for 1.18 s (Pearson correlation two-sided: $r = 0.86$, $p < 0.001$, $n = 15$, linear fit: $y = 1.18x$, $R^2 = 0.96$; Figure 17: B). It is to note that the data for the complex trials is not distributed normally (Kolmogorov-Smirnov: $D(15) = 0.30$, $p = 0.001$).

The delay duration however affected the mean fixation duration (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean duration of an object fixation: $F(1, 14) = 32.02$ $p < 0.001$, partial eta-square 0.696). In trials with a short delay the average fixation duration was 0.51 ± 0.043 s per fixation, while in trails with a longer delay the fixation duration was 0.44 ± 0.035 s per fixation (Figure 17: C). For every second a fixation lasted in a short delay trial a fixation on a long delay trial was 0.14 seconds longer (Pearson correlation two-sided: $r = 0.98$, $p < 0.001$, $n = 15$, linear fit: $y = 1.14x$, $R^2 > 0.999$) (Figure 17: D).

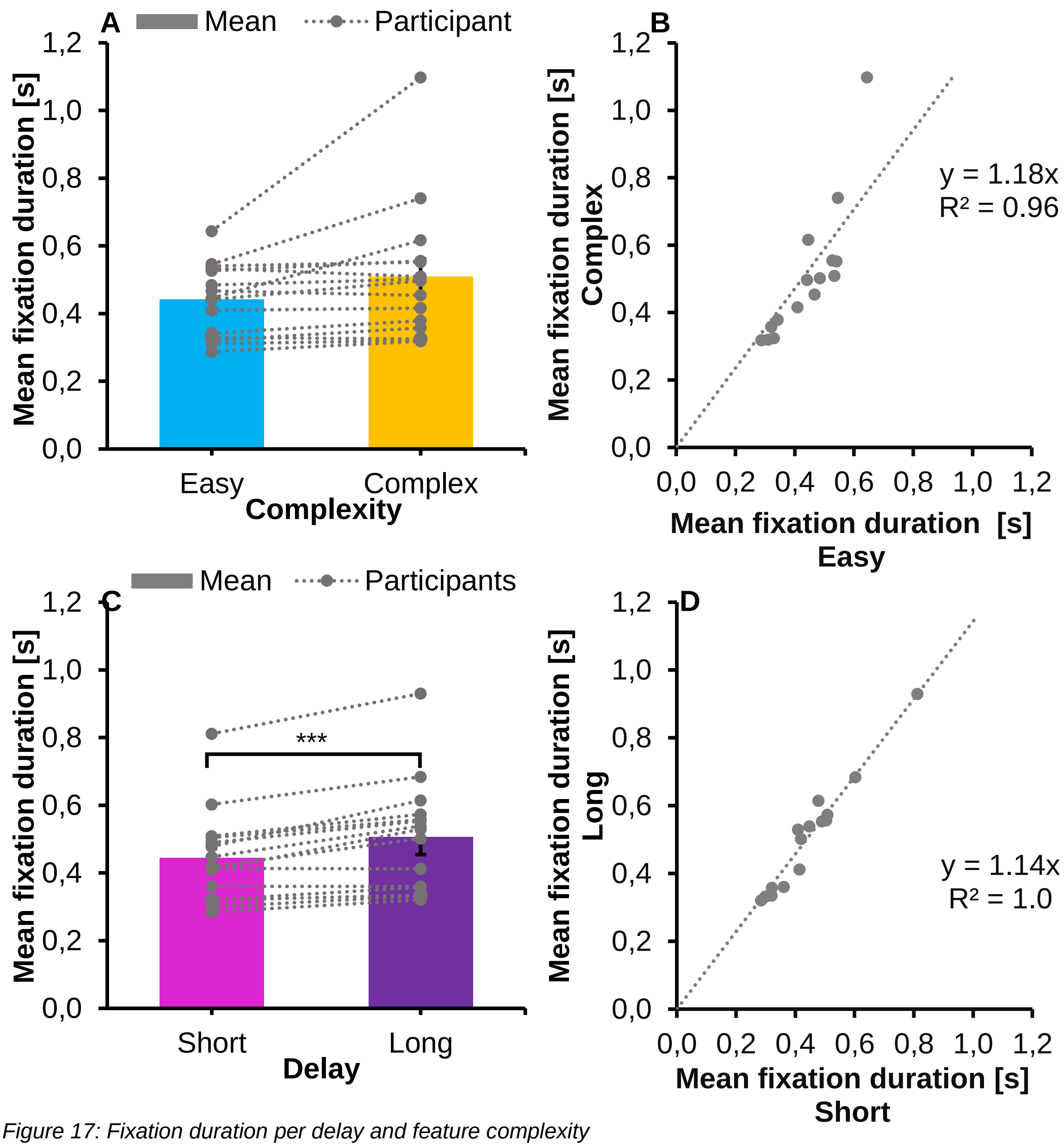

Figure 17: Fixation duration per delay and feature complexity
Mean duration of object fixations [s] plotted against the
A: Delay: short (pink) and long (purple) averaged over the participants (n = 15).
B: Correlation of mean fixation duration between easy and complex trials.
C: Complexity: Easy (blue) and complex (yellow) averaged over the participants (n = 15).
D: Correlation of mean fixation duration between short and long delays.

There was an interaction between Block and delay in duration of an object fixation (Multi-factorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean duration of an object fixation: $F(1, 14) = 15.92$, $p = 0.001$, partial eta-square 0.53). Participants mean duration of an object fixation was longer in Block 1 when the delay was long compared to the mean duration of an object fixation for trials with a short delay (Figure 18: A).

There was also an interaction between Block and feature regarding the duration of an object fixation in which participants tended to fixate complex objects longer in comparison to easy objects. This however changed in Block 2 were participants tended to fixate complex objects longer compared to easy objects (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean duration of an object fixation: $F(1, 14) = 5.32$, $p = 0.037$, partial eta-square 0.28; Figure 18: B).

The duration of an object fixation did not differ between sides (Paired two-sided t-test: $t(14) = 0.36$, $p = 0.72$).

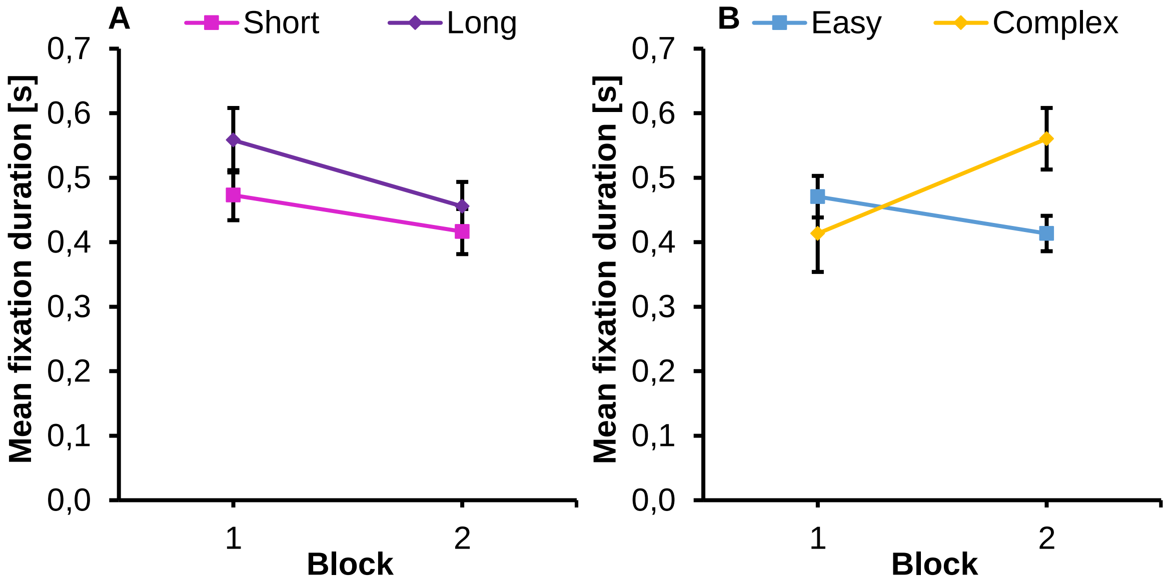

Figure 18: Interaction of delay and Block on the mean object Fixation duration
Mean object Fixation durations [s] plotted against the Block for
A: Short (pink) and long (purple) delays averaged over participant (n = 15).
B: Easy (yellow) and complex (blue) complexities averaged over participant (n = 15).

3.2 On a trial level

3.2.1 Number of fixations

The second parameter that can be analyzed is the number of fixations per trial. With an average of 90.16 ± 10.05 object fixations per trial and only 13.93 ± 1.47 non object fixations per trial within the screen per participant, the number of object fixations and non-object fixations per trial differ vastly, as seen before (Paired two-sided t-test: $t(14) = 1.41$, $p = 0.19$). It is to note that the mean number of object fixations per trials is not distributed normally (Kolmogorov-Smirnov: $D(15) = 0.25$, $p = 0.014$). Especially, the difference between the participants regarding the number of object fixations was noticeable (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean number of fixations per trial: $F(1, 14) = 80.43$, $p < 0.001$, partial eta-square 0.85). While participant 11 had the most object fixations with an average of 168.18 ± 9.85 fixations per trial, participant 8 had the fewest with an average of 48.43 ± 2.04 object fixations per trial (Figure 19; VIII. Supplementary: Figure 70).

To analyze the effects of Block, object complexity and delay further analysis was conducted. It is to note that several datasets are not distributed normally: the data for Block 1 with complex objects and both the long and the short delay, (Kolmogorov-Smirnov: $D(15) = 0.33$, $p < 0.001$; $D(15) = 0.23$, $p = 0.028$) as well as the data from Block 2 for complex objects and both long, and short delays (Kolmogorov-Smirnov: $D(15) = 0.27$, $p = 0.005$; $D(15) = 0.26$, $p = 0.007$) and lastly Block 2 for the easy objects and short delay condition (Kolmogorov-Smirnov: $D(15) = 0.24$, $p = 0.022$).

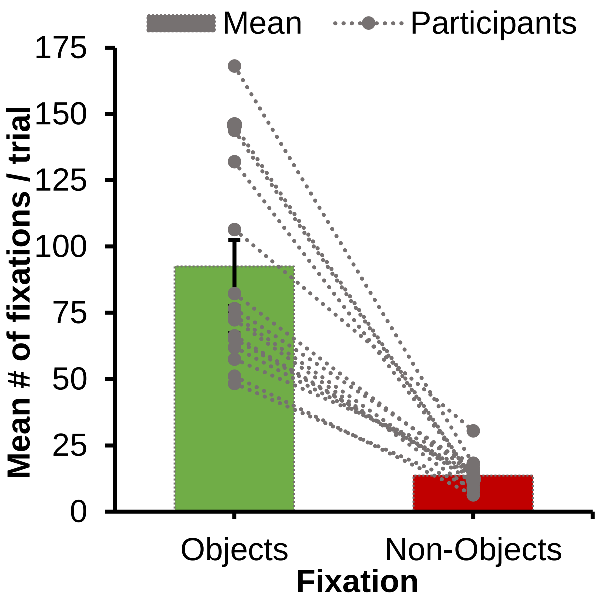

Figure 19: Number of Fixations per type of fixation
Number of fixations is plotted against the fixations type: Object (green) and non-object (red) averaged over participants (n = 15).

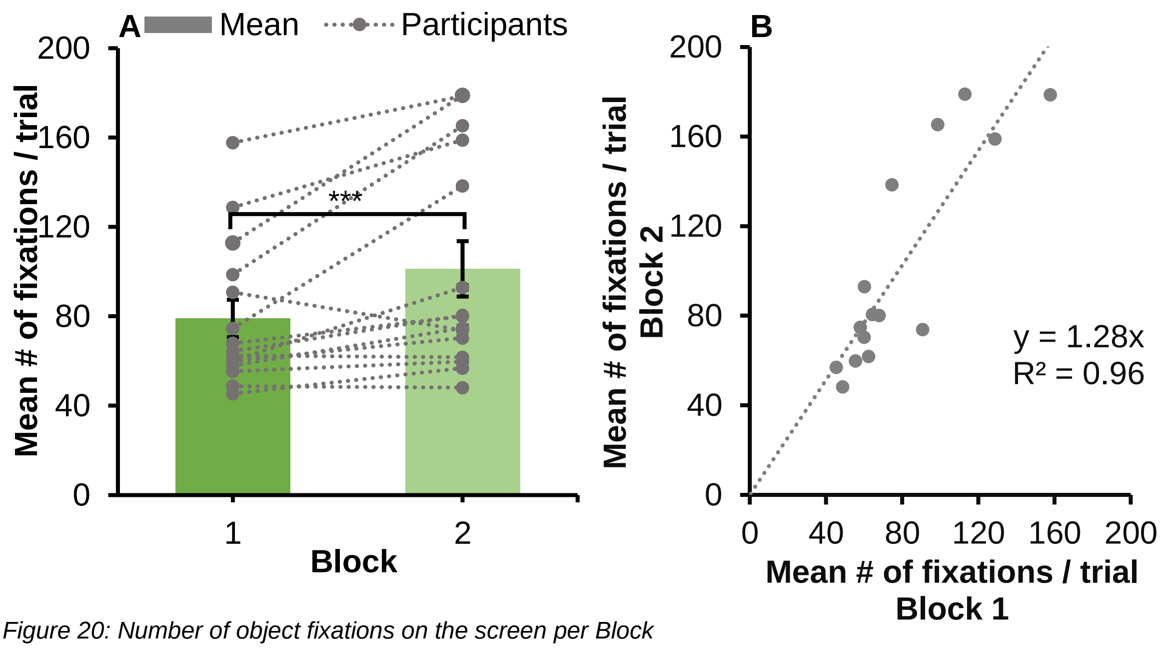

Figure 20: Number of object fixations on the screen per Block
A: Mean number of object fixations per trial averaged over the participants (n=15) as bars and for each participant as dots plotted against the Block.
B: Mean number of fixations per trial correlated for both Blocks (n = 15).

While the participants fixated the objects an average of 79.07 ± 8.38 times per trial in Block 1, the number of fixations per trial increased to an average of 101.26 ± 12.40 in Block 2 (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean number of fixations per trial: $F(1, 14) = 11.34$, $p = 0.005$, partial eta-square 0.45) (Figure 20: A). For every fixation done in Block 1 participants did 1.28 fixations in Block 2 (Pearson correlation: $r = 0.87$, $p < 0.001$, $n = 15$; Linear fit: $y = 1.28x$, $R^2 = 0.96$; Figure 20: B).

The number of object fixations per trial was less for the easy objects compared to complex objects (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean number of fixations per trial: $F(1, 14) = 20.01$, $p = 0.001$, partial eta-square 0.59). While participants fixated the objects an average of 77.29 ± 8.10 times per trial in trials with easy objects, the average number of object fixations increased to 103.04 ± 12.38 fixations per trial in trials with complex objects (Figure 21: A). For every fixation a participant did in the easy condition, 1.34 fixations were done in the complex condition (Pearson correlation: $r = 0.93$, $p < 0.001$, $n = 15$; Linear fit: $y = 1.34x$, $R^2 = 0.98$; Figure 21: B).

The delay duration had no significant effect on the number of object fixations (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean number of fixations per trial: $F(1, 14) = 2.58$, $p = 0.13$, partial eta-square 0.16).

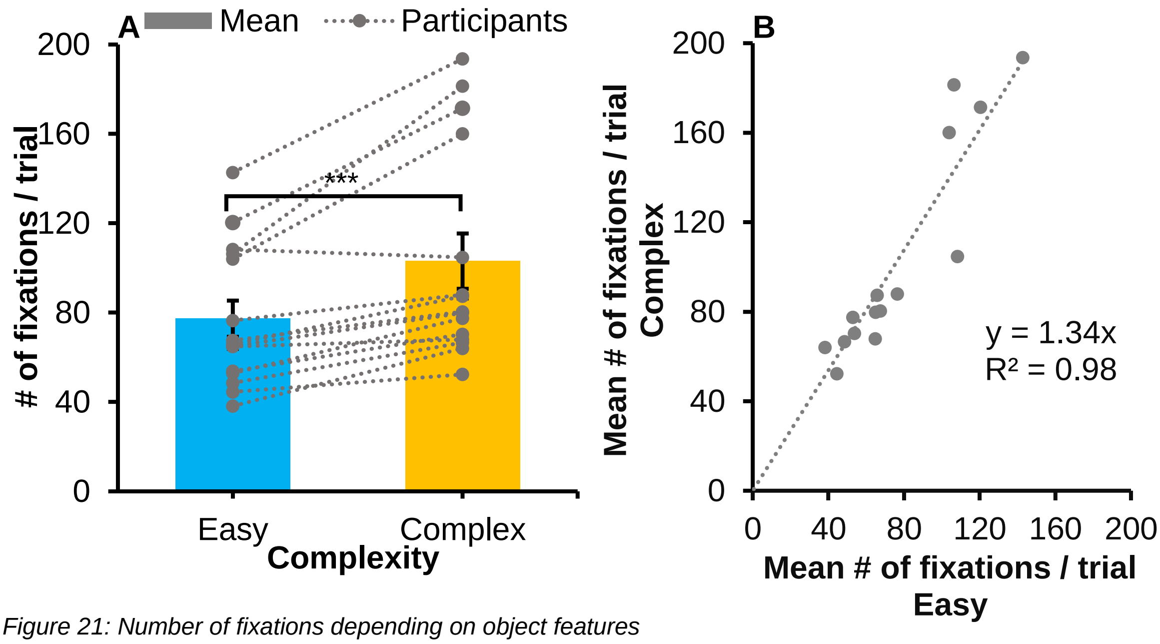

Figure 21: Number of fixations depending on object features
A: Mean number of fixations is plotted against the object feature: easy (yellow), complex (blue) averaged over the participants (n = 15).
B: Mean number of fixations per trial correlated for easy and complex trials (n = 15).

Additionally, there was an interaction between Block and complexity regarding the number of object fixations (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean number of fixations per trial: $F(1, 14) = 6.97$, $p = 0.019$, partial eta-square 0.33). For Block 2 there were more object fixations for complex trials compared to the number of object fixations in Block 2 for simple trials (Figure 22: A).

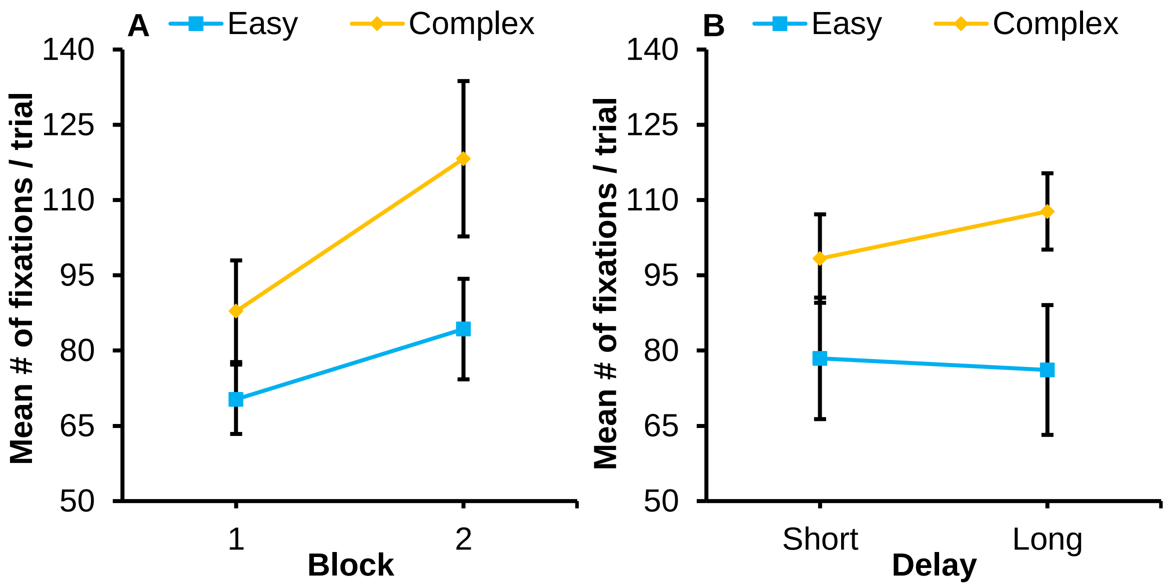

Figure 22: Number of object fixations per trial depending on feature complexity and Block or delay
A: Mean number of object fixations per trial plotted against the Blocks for both feature complexities: easy (blue) and complex (yellow) averaged over participants (n=15).
B: Mean number of object fixations plotted against the delay: short and long for both feature complexities: easy (blue) and complex (yellow) averaged over participant (n=15).

Furthermore, there was an interaction between delay and feature regarding the number of object fixations (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean number of fixations per trial: $F(1, 14) = 4.65$, $p = 0.049$, partial eta-square 0.25; Figure 22: B). In trials with complex objects the participants tended to fixate the objects more often especially in trials with a long delay compared to trials with short delay.

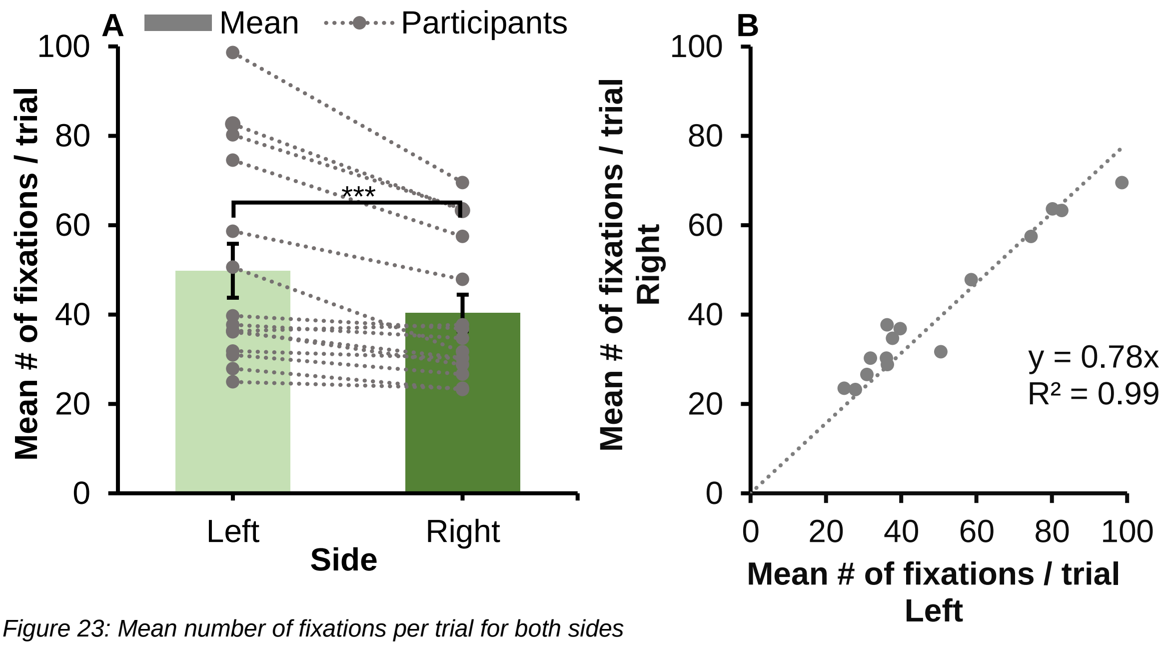

Figure 23: Mean number of fixations per trial for both sides
A: Mean number of fixations per trial plotted against the side averaged over all participants (n=15)).
B: Correlation of number of fixations on the left and number of fixations on the right side per trial over all participants (n = 15).

There were more fixations on the left compared to the right screen side (Paired two-sided t-test: $t(14) = 4.17$, $p = 0.001$). On the left side the participants fixated objects an average of 49.80 ± 4.08 times per trial while they fixated the objects on the right side only 40.36 ± 6.04 times per trial (Figure 23: A). It is to note that the number of fixations per trial is not distributed normally for either the left nor the right side (Kolmogorov-Smirnov: $D(15) = 0.27$, $p = 0.005$; $D(15) = 0.23$, $p = 0.028$). For every fixation participants did on the left side they did only 0.78 fixations on the right side (Pearson correlation: $r = 0.97$, $p < 0.001$, $n = 15$; Linear fit: $y = 0.78x$, $R^2 = 0.99$; Figure 23: B).

Further investigation has shown that the number of fixations per trial differed between the objects (Multifactorial repeated measurement ANOVA with independent factors side and object number and dependent factor mean number of fixations per trial, Greenhouse-Geisser corrected due to violation of sphericity: $F(3.32, 46.44) = 18.30$, $p < 0.001$, partial eta-square 0.57). Especially the last two objects where fixated less compared to the first ones (Figure 24).

Furthermore, there was an interaction between the side and the object number regarding the number of fixations per trial (Multifactorial repeated measurement ANOVA with independent factors side and object number and dependent factor mean number of fixations per trial, Greenhouse-Geisser corrected due to violation of sphericity: $F(3.06, 42.79) = 7.27$, $p < 0.001$, partial eta-square 0.34). On the left side the number of fixations per trial was larger for the first two objects compared to the corresponding objects on the right side (Bonferroni corrected two- sided paired t-tests: $t(14) = 6.16$, $p < 0.001$; $t(14) = 4.89$, $p < 0.001$).

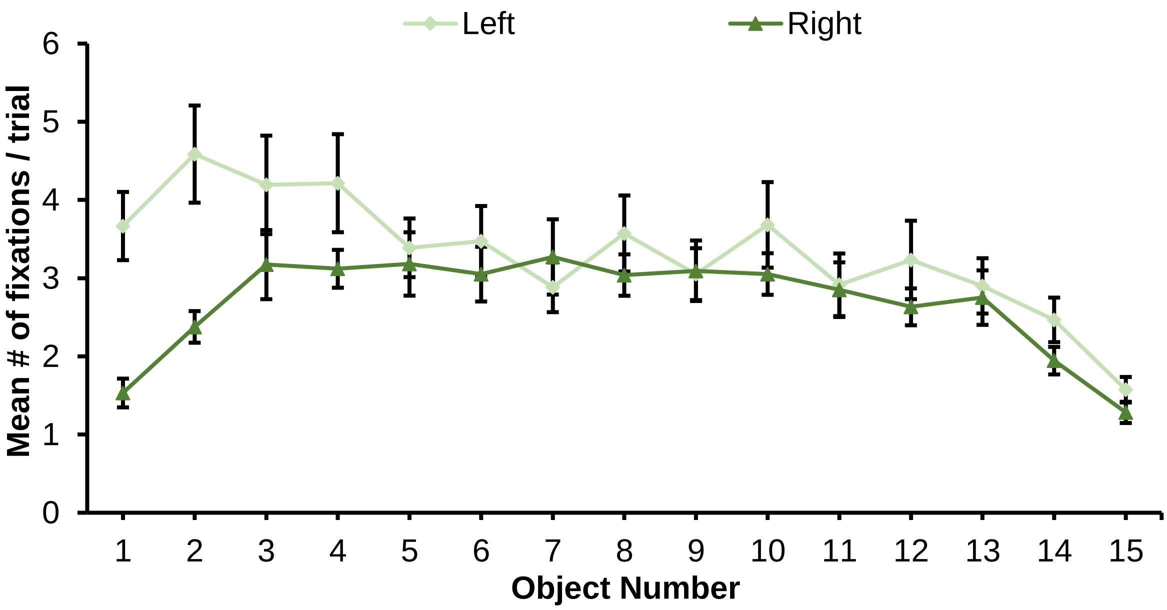

Figure 24: Mean number of fixations per trial for corresponding objects
Mean number of fixations per trial plotted against the object number (averaged over participant n= 15). Objects on the right side: dark green and object on the left side light green.

3.2.2 Total fixation time

Taking both the number of fixations as well as the fixation duration into consideration, we can look at the total fixation time per trial.

As seen before participants differed in their Total fixation time per trial (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean total fixation time per trial: $F(1, 14) = 170.06$, $p < 0.001$, partial eta-square 0.92) similar to findings from the fixation duration and the number of fixations. Participant 13 fixated objects the longest with an average Total fixation time of 59.44 ± 3.55 s per trial. On the contrary participant 10 had the shortest Total fixation time per trial with only 21.83 ± 1.10 s per trial on average (Figure 25). Over all participants the average Total fixation time per trial was 39.53 ± 3.02 s.

Different from the number of fixations per trial and the fixation duration, the mean total fixation time per trial did not differ between the Blocks (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean total fixation time per trial: $F(1, 14) = 1.26$, $p = 0.28$, partial eta-square 0.083). In Block 1 the average Total fixation time per trial was 37.85 ± 2.65 s in Block 1 and only 3 s more (40.85 s ± 3.84) in Block 2.

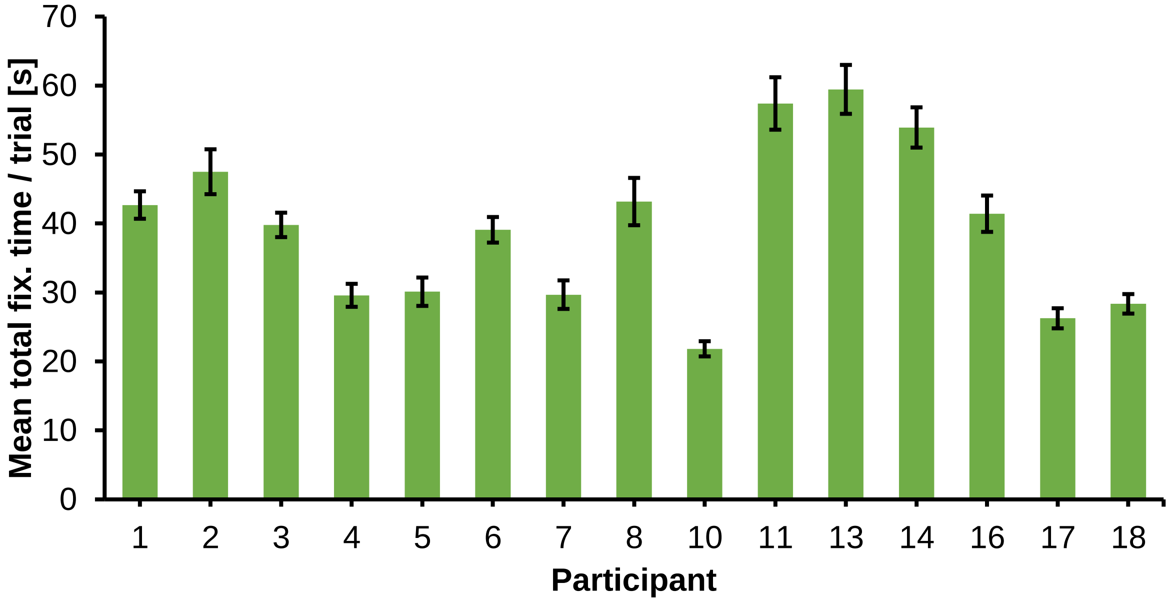

Figure 25: Mean total fixation time per trial for participants
Mean Total fixation time per trial [s] averaged over all trials (n=40) is plotted against the participant.

The delay duration increased the total fixation time per trial (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean total fixation time per trial: $F(1, 14) = 46.56$ $p < 0.001$, partial eta-square 0.77). In trials with a short delay the mean total fixation time per trial was 36.32 ± 3.22 s, while it was 42.37 ± 2.87 s in trials with a longer delay (Figure 26:A). For every second participants fixated in short trials the total fixation time per trial increased by 1.14 s in long trials (Pearson correlation: r = 0.96, $p < 0.001$, n = 15; Linear fit: y = 1.14x, $R^2 = 0.99$; Figure 26: B).

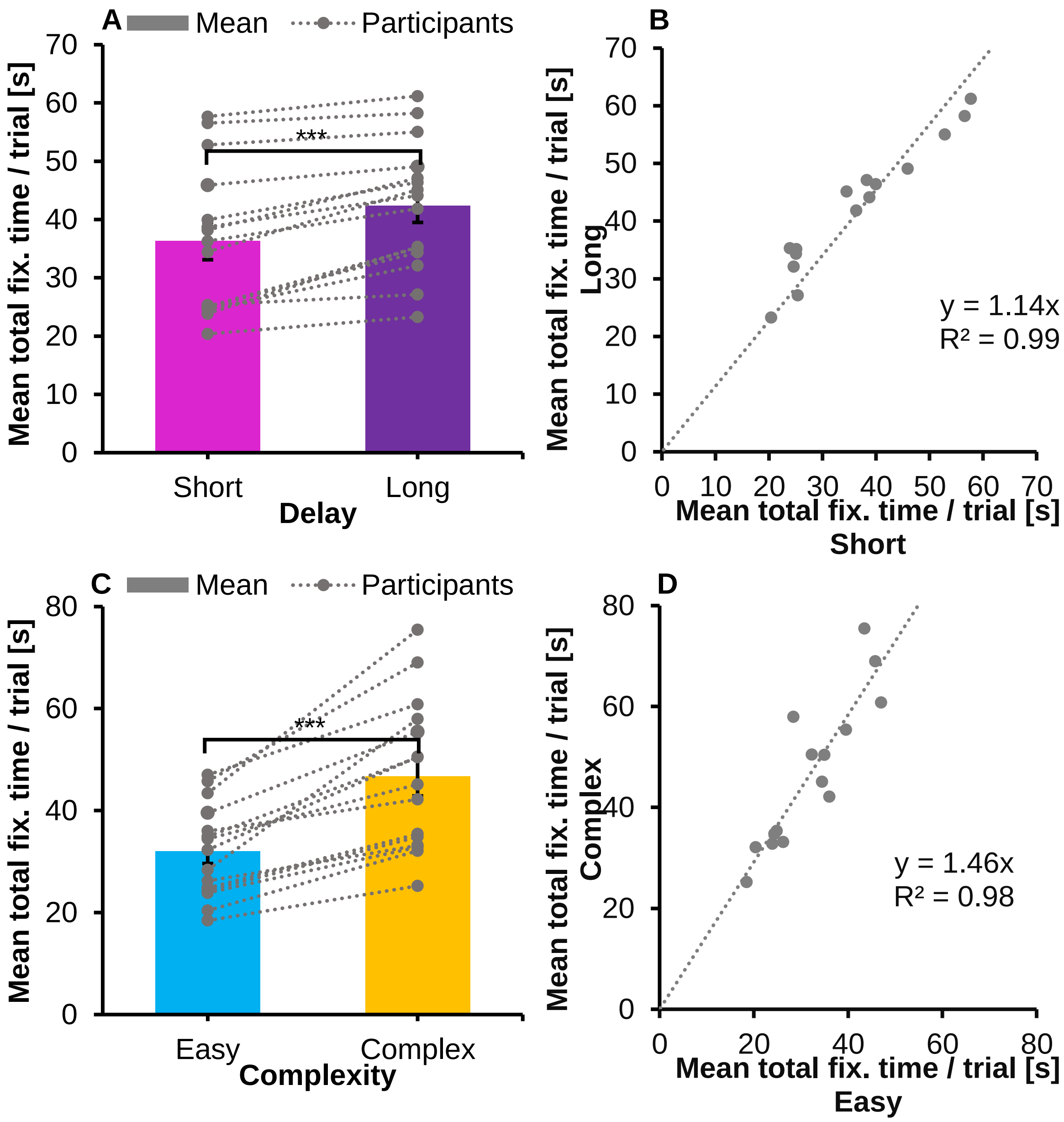

Figure 26: Mean Total fixation time per delay and feature
A: Mean Total fixation time per trial [s] plotted against the delay: short (pink) and long (purple) averaged over all participants (n = 15) as bars as well as for each participant as dots.
B: Mean number of fixations per trial correlated for short and complex trials (n = 15).
B: Mean Total fixation time per trial [s] plotted against the feature: easy (yellow) and complex (blue) averaged over all participants (n = 15) as bars as well as for each participant as dots.
D: Mean number of fixations per trial correlated for easy and complex trials (n = 15).

In addition to the delay, the complexity increased the total fixation time per trial as well (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean total fixation time per trial: $F(1, 14) = 50.73$ $p < 0.001$, partial eta-square 0.784). While participants fixated the objects for an average of 31.98 ± 2.36 s per trial in trials with easy object features, in trials with complex features the mean total fixation time was nearly 15 s more with 46.71 ± 3.84 s per trial (Figure 26:

C). For every second participants fixated easy objects per trial, the number of seconds complex objects were fixated per trial increased by nearly 1.46 s (Pearson correlation: r = 0.89, $p < 0.001$, n = 15; Linear fit: y = 1.46x, $R^2 = 0.99$) (Figure 26: D).

There was an interaction between Block and delay regarding the total fixation time per trial (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean total fixation time per trial: $F(1, 14) = 21.62$, $p < 0.001$, partial eta-square 0.61), in which participants fixated shorter when the delay was short especially in Block 1 (Figure 27 A).

Furthermore, there was in interaction between delay and feature regarding the mean total fixation time in which participants tended to fixate longer in trials with complex features and a longer delay compared to a shorter delay (Multifactorial repeated measurement ANOVA with independent factors: Block, feature, delay and dependent factor: total fixation time for n = 15 participants: $F(1, 14) = 10.06$, $p = 0.007$, partial eta-square 0.418; Figure 27 B).

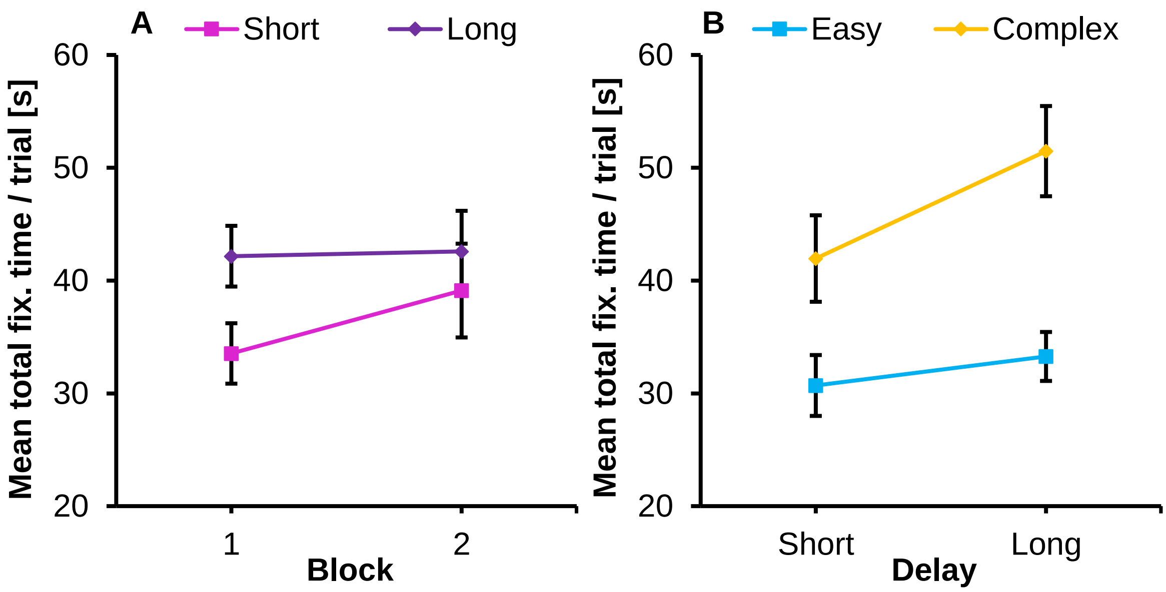

Figure 27: Total fixation time depending on delay and feature
Mean Total fixation time per trial [s] plotted against the
A: Block for the delay: short (pink) and long (purple) averaged over the participants (n = 15).
B: Delay for feature complexities: easy (yellow) and complex (blue) averaged over the participants (n = 15).

Corresponding with our findings regarding the number of fixations depending on the screen side we can see that the total fixation time per trial is longer on the left compared to the right side (two-sided t-test repeated measurement: $t(14) = 4.92$, $p < 0.001$). On the left screen side the mean total fixation time per trial was 21.46 ± 1.76 s while it was 17.88 ± 1.30 s on the right side (Figure 28: A). For every second fixated on the left side of the screen only 0.82 s were fixated on the right side per trial (Pearson correlation: r = 0.93, $p < 0.001$, n = 15; Linear fit: y = 0.82x, $R^2 = 0.99$; Figure 28: B).

Going more into detail, one can see that there were differences in how long an object was fixated per trial (Multifactorial repeated measurement ANOVA with independent factors side and object number and dependent factor mean number of fixations per trial, Greenhouse-Geisser corrected due to violation of sphericity: $F(4.244, 69{,}423) = 11.89$, $p < 0.001$, partial eta-square 0.46).

Furthermore, there was an interaction between the side and object number regarding the total fixation time (Multifactorial repeated measurement ANOVA with independent factors side and object number and dependent factor mean number of fixations per trial, Greenhouse-Geisser corrected due to violation of sphericity: $F(2.68, 37.45) = 6.03$, $p < 0.001$, partial eta-square 0.30). This results mainly from the difference between the first two and the last objects (Bonferroni corrected two- sided paired t-tests: $t(14) = 5.43$, $p = 0.001$; $t(14) = 4.50$, $p = 0.008$; $t(14) = 3.57$, $p = 0.046$; Figure 29).

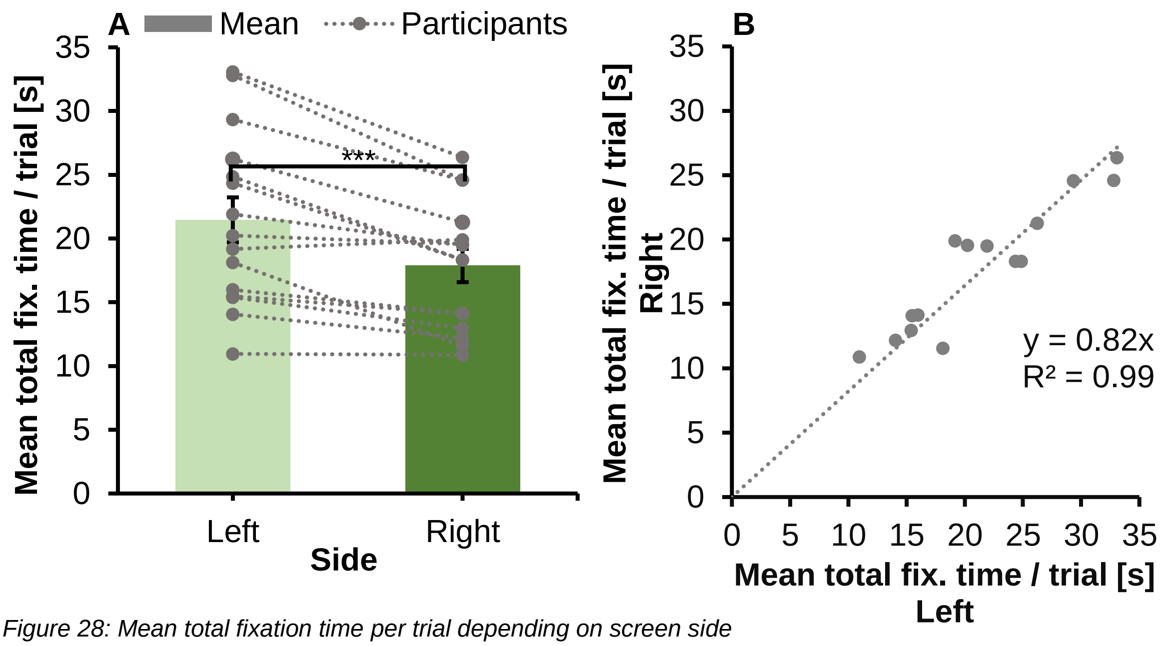

Figure 28: Mean total fixation time per trial depending on screen side
A: Mean total fixation time per trial [s] plotted against the screen side: right (light green) and left (dark green) averaged over all participants (n = 15) as bars as well as for each participant as dots.
B: Correlation between the mean total fixation time per trial between the left and the right side plotted for all participants (n = 15).

Similar findings regarding the difference between objects and sides were observed in the number of fixations per trial.

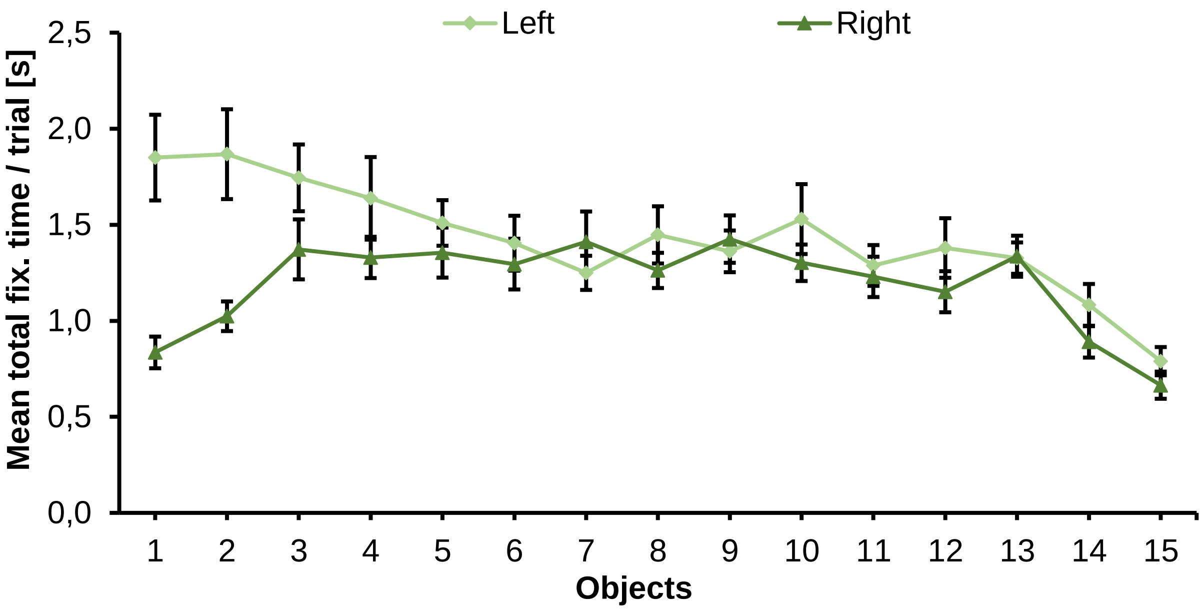

Figure 29: Total fixation time per object and side
Overall object the mean total fixation time [s] plotted against the objects and the side (left: light green and right: dark green) averaged over the participants (n=15).

3.3 Number of hemifield switches

Now we will look at the number of hemifield switches per trial. Participants switched an average of 9.91 ± 0.98 times per trial (Figure 30). The average number of switches per trial varied between 4.75 ± 0.33 for participant 7 to 16.48 ± 0.94 switches for participant 3 (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean number of switches per trial: $F(1, 14) = 102.02$, $p < 0.001$, partial eta-square 0.88).

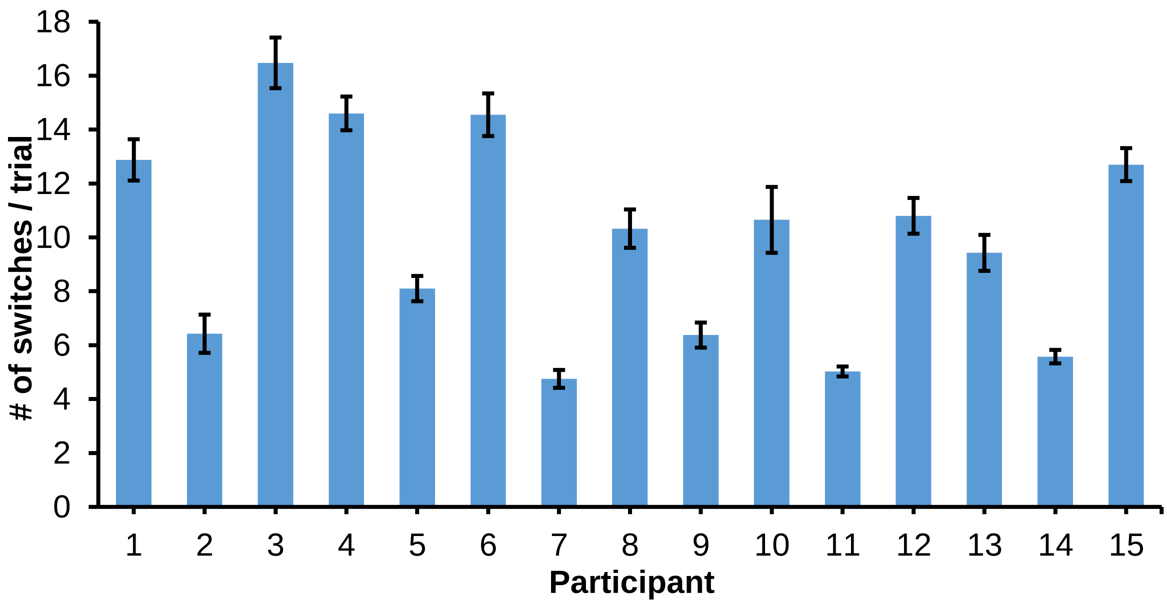

Figure 30: Mean number of switches per trial for participants
Mean number of switches per trial averaged over trials (n = 40) plotted against the participant.

Block 2 had fewer switches compared to Block 1 (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean number of switches per trial: $F(1, 14) = 36.59$, $p < 0.001$, partial eta-square 0.72). While the participants switched on average 12.29 ± 1.31 times per trial in Block 1, the number of switches decreased to 7.53 ± 0.72 in Block 2 (Figure 31: A). For every switch done in Block 1 within a trial 0.59 switches were done in Block 2 (Pearson correlation coefficient: $r = 0.89$, $p < 0.001$, $n = 15$, Linear fit: $y = 0.59x$, $R^2 = 0.96$; Figure 31: B).

With the forced reduction of switches to 75% in Block 2 we expected an average of 9.22 switches per trial, however we observed an average of 7.53 switches per trial which equals a reduction to 61.23% compared to Block 1. The number of switches per trial was different from the expected number of switches per trial, which was given by the reduction to 75% of Block 1 (One- sample t-test: $t(14) = 2.35$, $p = 0.034$).

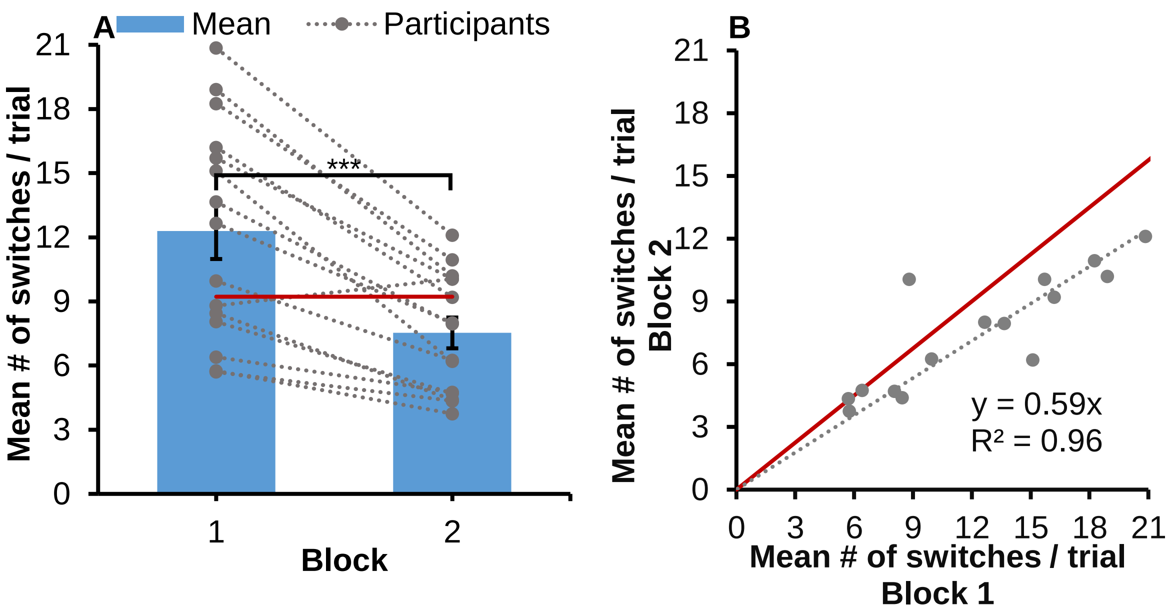

Figure 31: Number of switches for Blocks
A:Number of switches per trial plotted against the Block averaged over all participants (n = 15) as bars as well as for each participant as dots. The red line visualizes the reduction to 75% of the number of switches made in Block 1.
B: Number of switches per trial in Block 2 plotted against number of switches per trial in Block 1 for n = 15 participants. The red line visualizes the reduction to 75% of the number of switches made in Block 1.

It can be seen that in trials with easy objects, participants needed less switches compared to trials with complex objects (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean number of switches per trial: $F(1, 14) = 22.13$, $p < 0.001$, partial eta-square 0.61). In trials with easy features participants needed on average 8.90 ± 0.95 switches, while they needed 10.92 ± 1.05 switches. This makes on average two switches more with in trials with complex objects compared to easy ones (Figure 32: A). For every switch done in an easy trial, participants did 1.2 switches in a complex trial (Pearson correlation coefficient: $r = 0.91$, $p < 0.001$, $n = 15$, Linear fit: $y = 1.20x$, $R^2 = 0.98$).

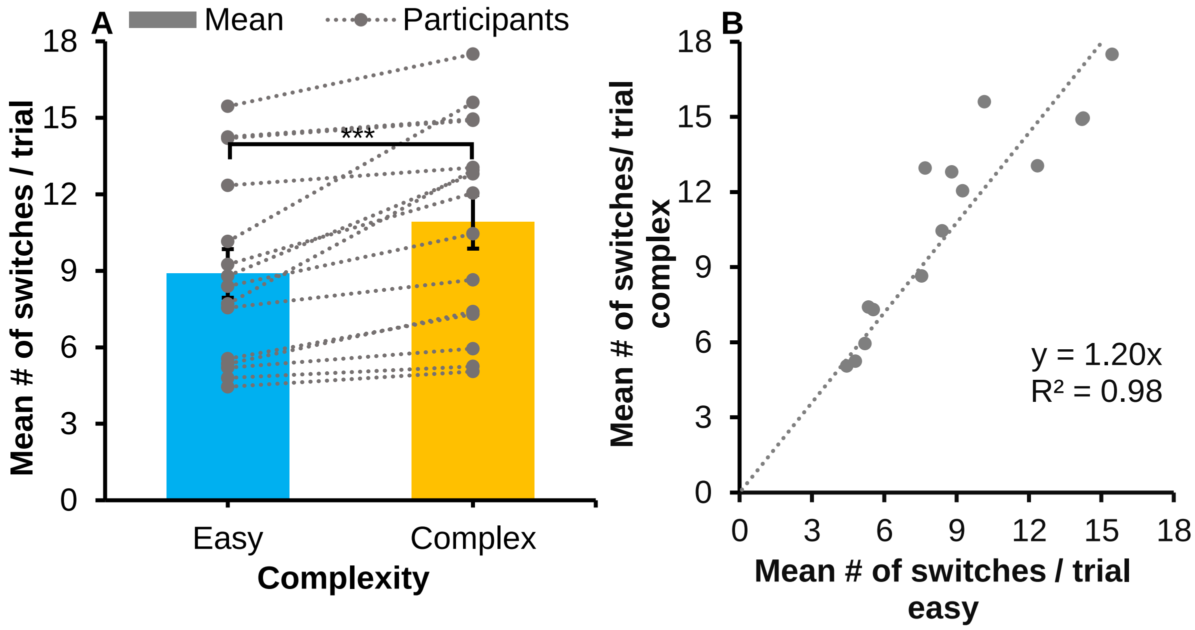

Figure 32: Mean number of switches per trial for feature complexity
A: Mean number of switches per trial is plotted against the feature complexity easy (yellow), complex (blue) averaged over participant (n = 15)
B: Correlation of the number of switches per trial between easy and complex objects averaged over all participants (n = 15).

There was an interaction between feature complexity and Block, in which participants needed more switches in Block 1 especially in trials with complex feature complexities compared to trials with easy features (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean number of switches per trial: $F(1, 14) = 6.66$, $p = 0.022$, partial eta-square 0.32; Figure 33).

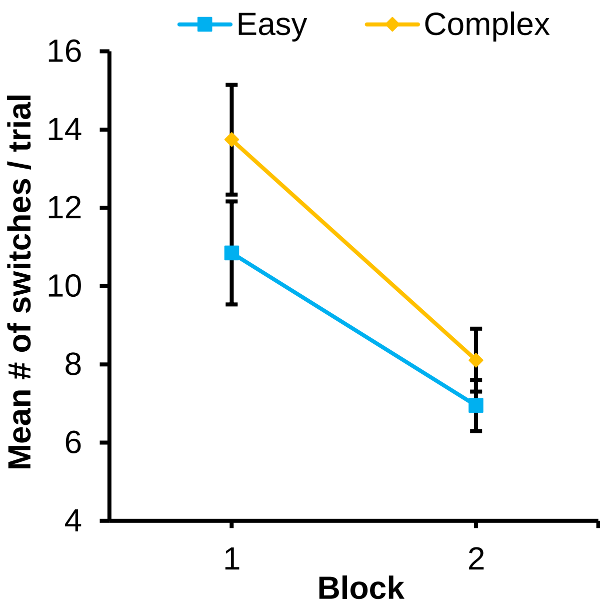

Figure 33: Interaction of Block and complexity in mean number of switches per trial
Mean number of switches per trial is plotted against the Block for both complexities: easy (yellow), complex (blue) average over participants (n = 15).

By looking at the switches and the sides it could be seen that 87 trials from 600 began on the right side of the screen, even though objects where visible only on the left. Those mis-fixations will be excluded in the following.

3.4 On a switch level

Similar to Pomplun et al. (2001), the successive fixations were also examined within a hemifield switch.

3.4.1 Number of fixations

The participants differed in the overall number of fixations per switch (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean number of fixations per switch: $F(1, 14) = 32.34$, $p < 0.001$, partial eta-square 0.70). The mean number of fixations per switch ranged from 4.52 ± 0.22 fixations per switch for participant 4 to 30.33 ± 2.00 fixations per switch for participant 13 (

Figure 34). The average number of fixations per switch was 12.86 ± 2.29.

The overall number of fixations per switch was smaller in Block 1 compared to Block 2 (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean number of fixations per switch: $F(1, 14) = 18.84$, $p = 0.001$, partial eta-square 0.57). While in Block 1 there were an average of 8.35 ± 1.52 fixations per switch, in Block 2 participants did on average 16.57 ± 3.03 fixations per switch (Figure 35: A). For every fixation per switch done in Block 1 1.89 fixations were done in Block 2 (Pearson coefficient: $r = 0.84$, $p < 0.001$, $n = 15$, Linear fit: $y = 1.89x$, $R^2 = 0.90$; Figure 35: B).

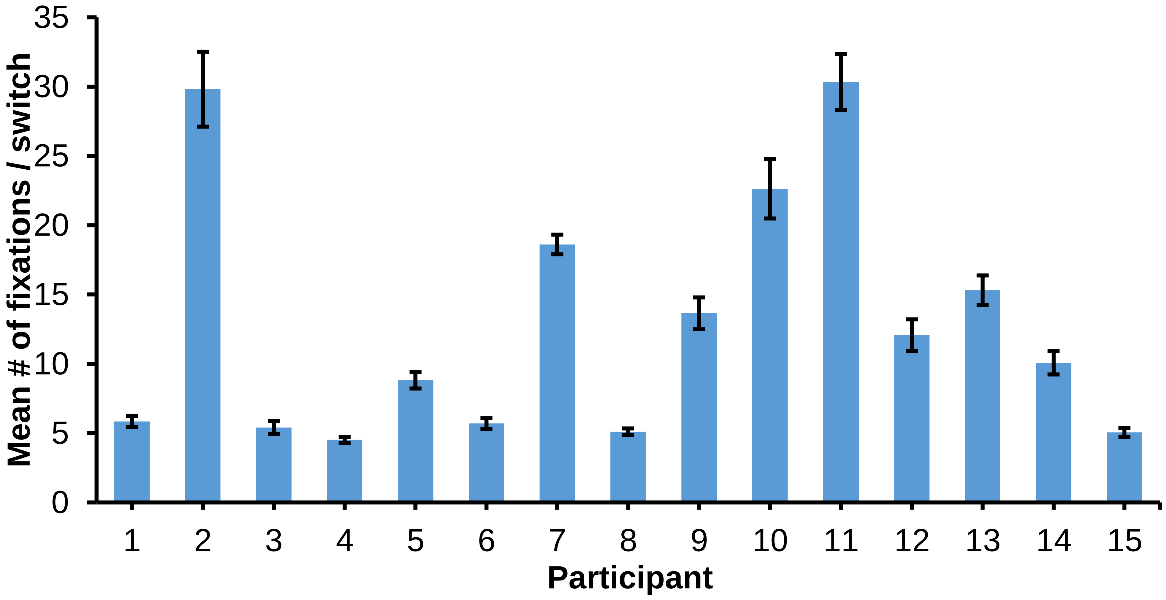

Figure 34: Mean number of fixations per switch for participants
Mean number of fixations per switch plotted against the participant for overall fixations averaged over the trials (n = 40).

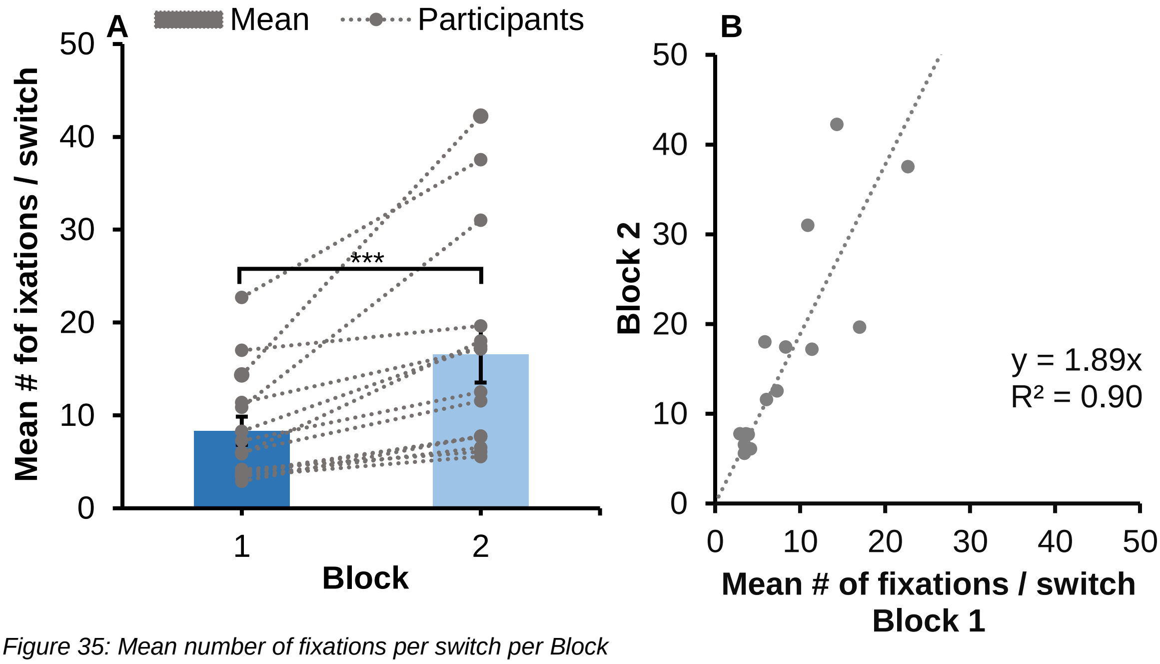

Figure 35: Mean number of fixations per switch per Block
A: Mean number of fixations per switch plotted against the Block averaged over all participants (n = 15) as bars as well as for each participant as dots.
B: Mean number of fixations per switch in Block 2 plotted against number of switches per trial in Block 1 for n = 15 participants.

Neither the feature complexity nor the delay had an effect on the mean number of fixations per switch (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean number of fixations per switch: $F(1, 14) = 2.53$, $p = 0.13$, partial eta-square 0.15; $F(1, 14) = 0.41$, $p = 0.41$, partial eta-square 0.048).

However, there was an interaction between Block and complexity (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean number of fixations per switch: $F(1, 14) = 6.11$, $p = 0.027$, partial eta-square 0.30). In Block 2 there were more fixations for complex trials compared to easy trials (Figure 36).

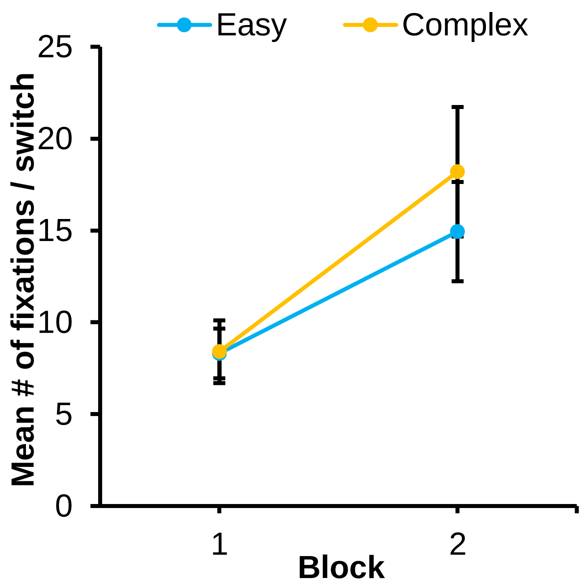

Figure 36: Interaction of Block and complexity in mean number of fixations per switch
Mean number of fixations per switch is plotted against the Block for both complexities: easy (yellow), complex (blue) average over participants (n = 15).

With a mean of 11.78 ± 2.12 fixations per switch on the left and only 10.67 ± 1.84 there were more fixations on the left compared to the right (two-sided paired t-test: $t(14) = 2.73$, $p = 0.016$). For every fixation made on the left side per switch only 0.89 fixations were made on the right side (Pearson coefficient: $r = 0.99$, $p < 0.001$, n=15, Linear fit: $y = 1.89x$, $R^2 = 0.99$).

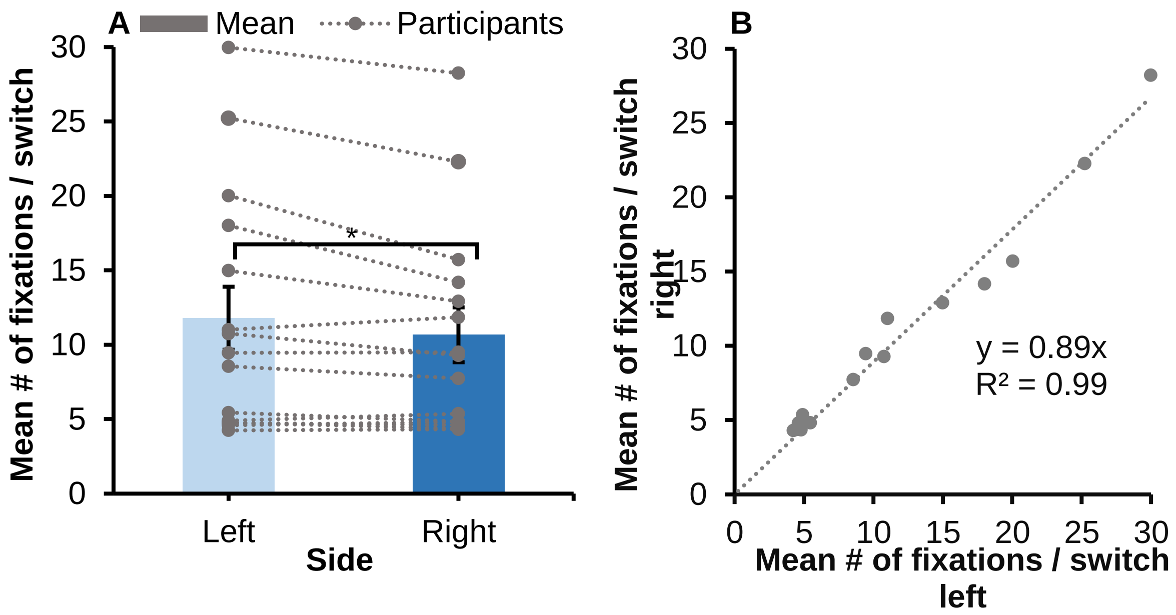

Figure 37: Mean number of fixations per switch depending on screen side
A: Mean number of fixations per switch plotted against the screen side: right (light green) and left (dark green) averaged over all participants (n = 15) as well as for each participant (grey).
B: Correlation between the mean number of fixations per switch between the left and the right side plotted for all participants (n = 15).

3.4.2 Total fixation time

The average total fixation time and therefore also the average processing time was 5.19 ± 0.74 s. The mean total fixation time per switch differed vastly between the participants (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean total fixation time per switch: $F(1, 14) = 48.42$, $p < 0.001$, partial eta-square 0.78) and ranged from 2.07 ± 0.09 s in participant 4 to 12.47 ± 0.85 s in participant 11 (Figure 38).

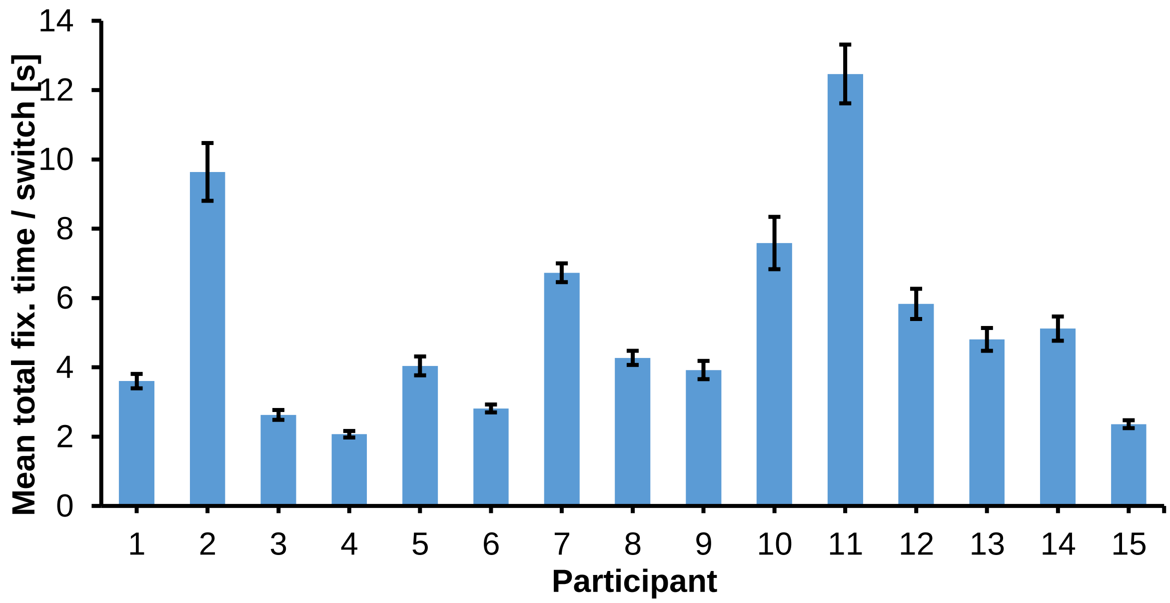

Figure 38: Mean total fixation time per switch for participants
Mean total fixation time [s] per switch plotted against the participant for overall fixations averaged over the trials (n = 40).

For further investigation it is to note that the mean total fixation time was not distributed normally for: Block 1, both the easy and the long delay in the complex condition and the short delay in the easy condition (Kolmogorov-Smirnov: D(15) = 0.26, p = 0.009; D(15) = 0.24, p = 0.022; D(15) = 0.31, p < 0.001). In Block 2 the data from both delays in the complex condition was non-normally distributed (Kolmogorov-Smirnov: D(15) = 0.25, p = 0.014; D(15) = 0.28, p = 0.002).

In Block 2 participants fixated longer per switch compared to Block 1 (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean total fixation time per switch: F(1, 14) = 19.39, p = 0.001, partial eta-square 0.581). While the average total fixation time per switch was 3.68 ± 0.52 s in Block 1, it increased by 75.5% to an average of 6.46 ± 1.00 s in Block 2 (Figure 39: A). Total fixation time per switch increased by 1.72 s in Block 2 for every second needed in Block 1 (Pearson coefficient: D(15) = 0.83, p < 0.001, Linear fit: y = 1.72, R^2= 0.92; Figure 39: B). It is to note, that the total fixation time per switch is not distributed normally in Block 1 (Kolmogorov-Smirnov: D(15) = 0.27, p = 0.005).

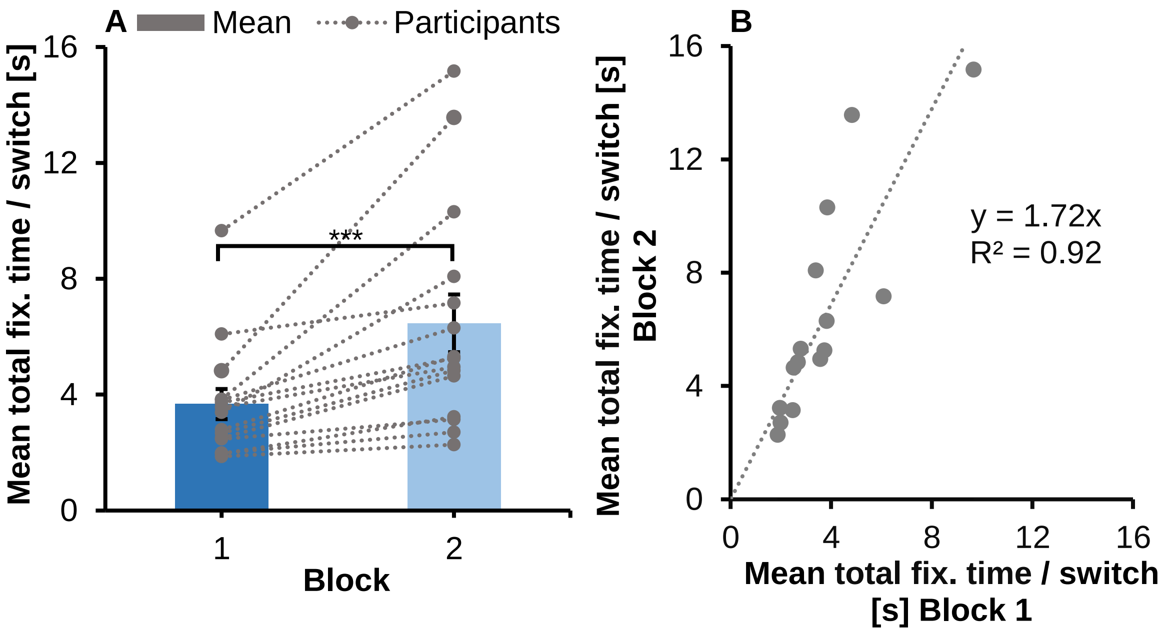

Figure 39 Mean total fixation time per switch for both Blocks
A: Mean total fixation time per switch [s] plotted against the Block averaged over all participants (n = 15) as bars as well as for each participant as dots.
B: Mean total fixation time per switch [s] in Block 2 plotted against number of switches per trial in Block 1 for n = 15 participants.

The feature complexity increased the mean total fixation time per switch (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean total fixation time per switch: F(1, 14) = 6.70, p = 0.019, partial eta-square 0.33). In trials with easy objects participants fixated on average 4.54 ± 0.60 s per switch, while they fixated more than one second longer (5.59 ± 0.88 s) in trials with complex objects (Figure 40: A). For every second easy objects were fixated per switch participants fixated complex objects 1.26 s (Pearson coefficient: D(15) = 0.92, p < 0.001, Linear fit: y = 1.26, R^2 = 0.96; Figure 40: B). It is to note that the mean total fixation time per switch for the easy trials was not distributed normally (Kolmogorov-Smirnov: D(15) = 0.24, p = 0.020).

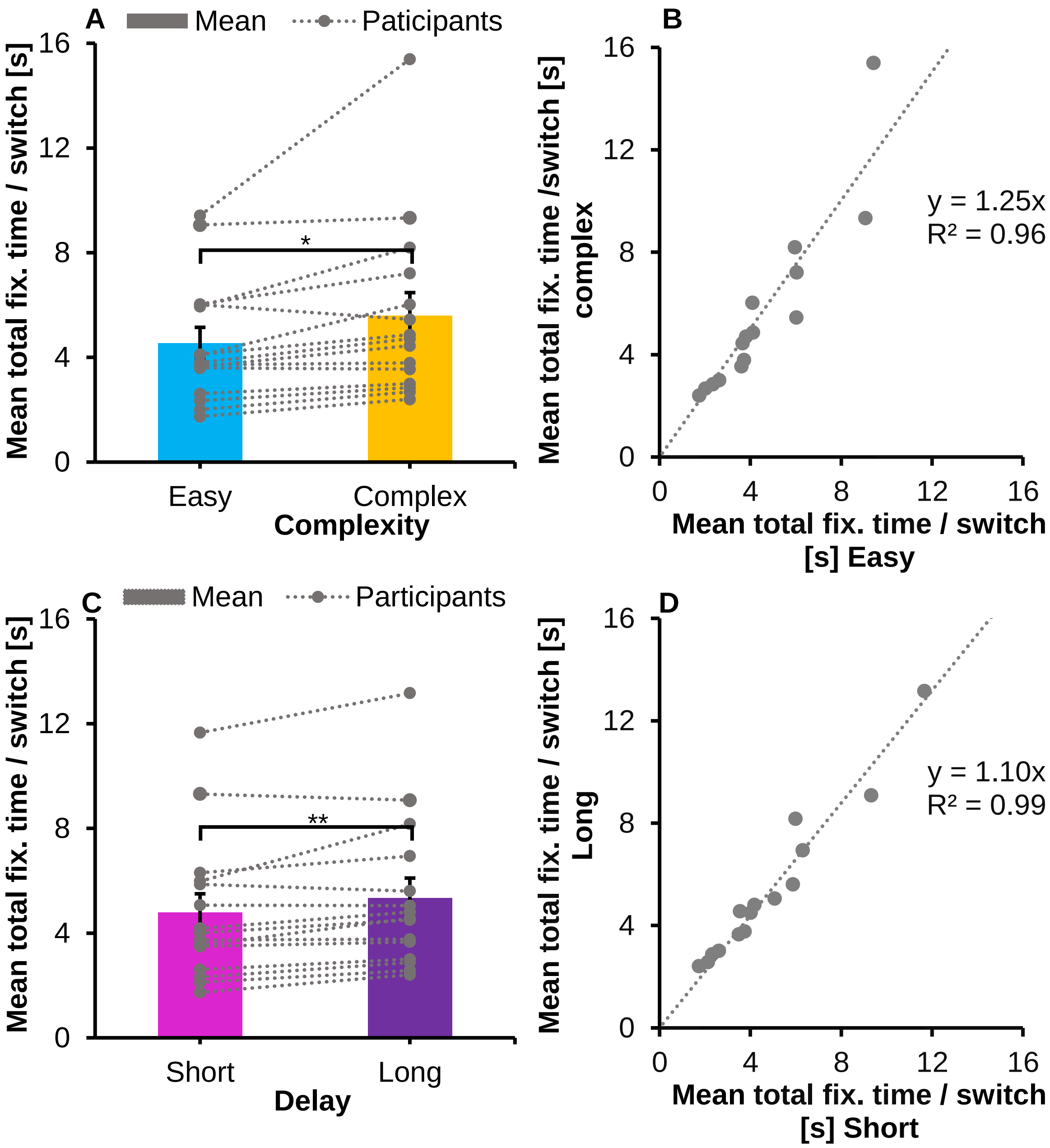

Figure 40: Mean total fixation time per switch for delay and feature
A: Mean total fixation time per switch [s] plotted against the feature: easy (yellow) and complex (blue) averaged over all participants (n = 15) as bars as well as for each participant as dots.).
B: Mean total fixation time per switch [s] correlated for easy and complex trials (n = 15).
C: Mean total fixation time per switch [s] plotted against the delay: short (pink) and long (purple) averaged over all participants (n = 15) as bars as well as for each participant as dots.
D: Mean number of fixations per switch correlated for short and complex trials (n = 15).

With a shorter delay present the total fixation time per switch was shorter compared to switches with a longer delay (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean total fixation time per switch: F(1, 14) = 10.80, p = 0.005, partial eta-square 0.44). The mean total fixation time per switch in short delay conditions was 4.79 ± 0.71 s, while the duration was on average 5.34 ± 0.76 s in the long delay condition (Figure 40: C). For every second per trial that participants fixated in short delay conditions, they fixated 0.1 seconds longer in the long delay condition (Pearson coefficient: D(15) = 0.98, $p < 0.001$, Linear fit: y = 1.10, R^2= 0.99; Figure 40: D).

There was an interaction between Block and complexity (Multifactorial repeated measurement ANOVA with independent factors: Block, complexity, delay and dependent factor: mean total fixation time per switch: F(1, 14) = 5.93, p = 0.029, partial eta-square 0.30). The total fixation time per switch was longer in Block 2 for complex trials compared to easy trials (Figure 41).

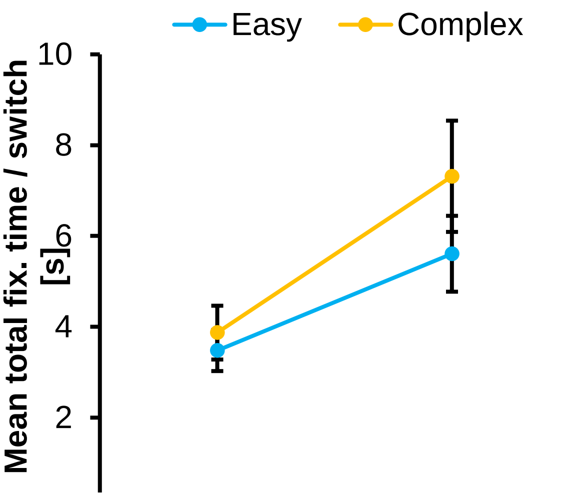

Figure 41: Mean total fixation time per switch interaction Block and complexity
Mean Total fixation time per switch [s] is plotted against the Block for both complexities: easy (yellow), complex (blue) average over participants (n = 15).

On the left side the total fixation time per switch was slightly longer with 4.85 ± 0.70 s compared to the right side with 4.52 ± 0.64 s (two-sided paired t-test t(14) = 2.37, p = 0.033) (Figure 42: A). For every second participants fixated per switch on the left side, they fixated only 0.92 s on the right side (Pearson coefficient: D(15) = 0.98, $p < 0.001$, Linear fit: y = 0.92, R^2= 0.99; Figure 42: B).

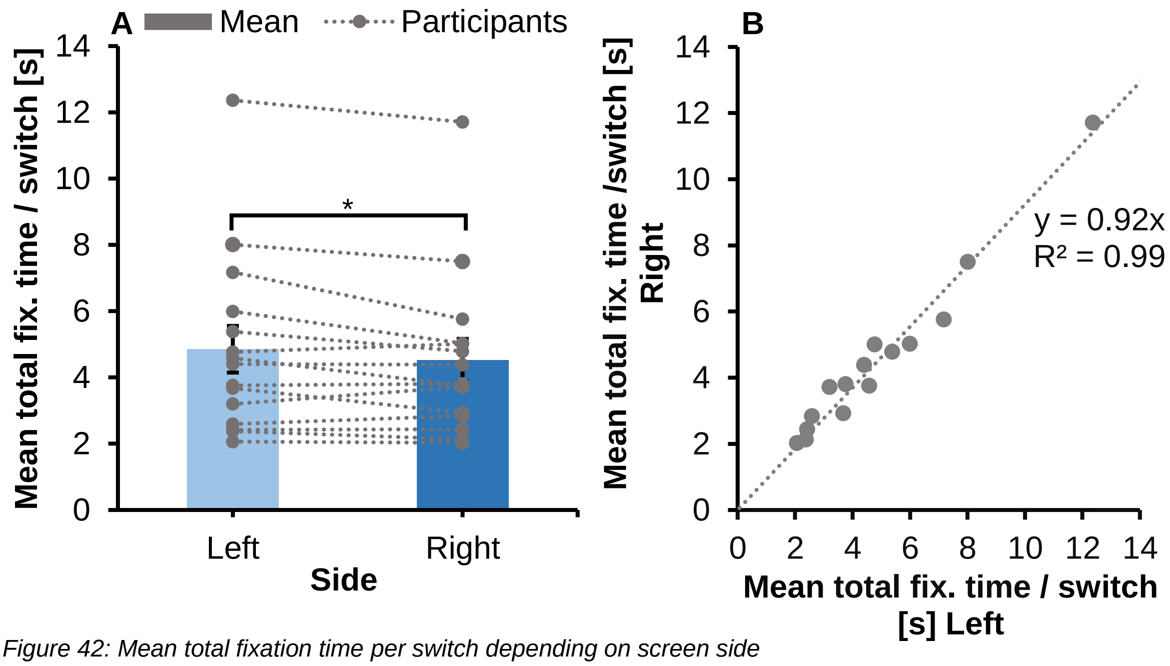

Figure 42: Mean total fixation time per switch depending on screen side
A: Mean total fixation time per switch [s] plotted against the screen side: right (light blue) and left (dark blue) averaged over all participants (n = 15) as bars as well as for each participant as dots.
B: Correlation between the mean total fixation time per switch [s] between the left and the right side plotted for all participants (n = 15).

3.5 Including the classification as chunk and revisit

3.5.1 Chunk size and number of revisits

As a short reminder: fixations within a switch can be classified either as *chunk*: the first fixation of an object within a switch or as a *revisit*: a refixation of an already visited object within one switch. The average number of chunk fixations per switch will be called chunk size in the following, while the average number of fixations per switch classified as revisit will be called number of revisits.

With a mean chunk size of 4.96 ± 0.60 and a mean number of revisits of 7.90 ± 1.78 those two parameters were different from another (Paired two-sided t-test: t (14) = -2.20, p = 0.045; Figure 43: A). With an increase in chunk size the number of revisits increased exponentially (Pearson coefficient: D(15) = 0.81, p < 0.001; Power fit: $y = 0.475x^{1.65}$, R^2 = 0.80; Figure 43: B).

Chunk size differed between the participants from an average of 2.48 ± 0.08 objects for participant 3 to an average of 8.98 ± 0.27 objects per chunk for participant 7 (Multifactorial repeated measurement ANOVA with independent factors Block, feature, delay and dependent factor mean chunk size: F(1, 14) = 71.55, p < 0.001, partial eta-square 0.836). For a detailed visualization see chapter VIII. Supplementary: Figure 71.

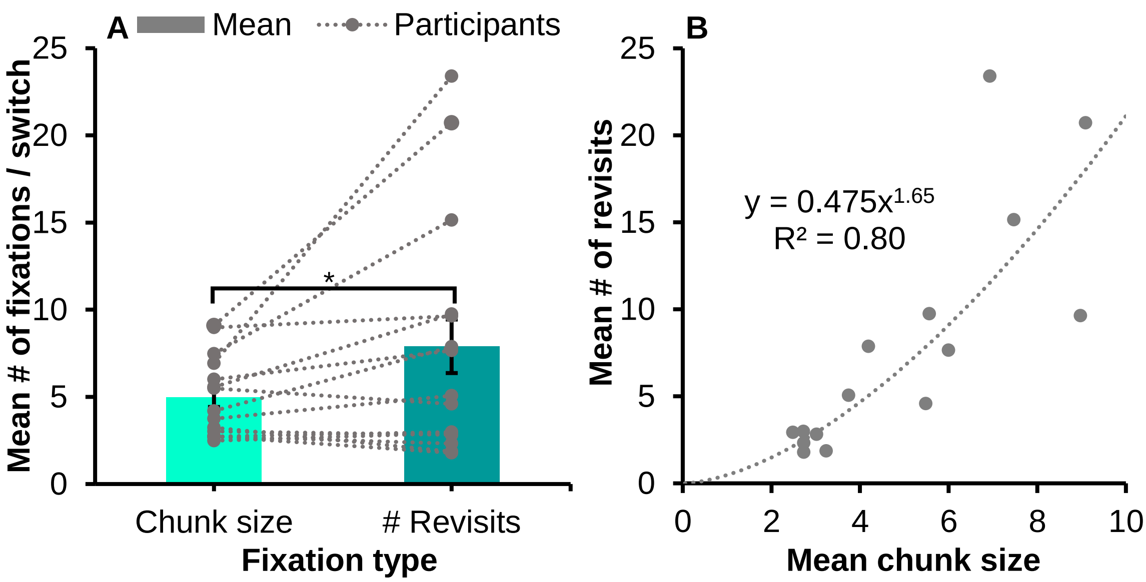

Figure 43: Mean chunk size and number of revisits
A: Mean number of fixations per switch is plotted against the fixation type: chunk (turquoise) and revisit (petrol) averaged over trial (n = 40) for each participant as dots and over all participants (n = 15) as bars.
B: Mean number of revisits plotted against the mean chunk size.

The number of revisits per switch also differed between participants (Multifactorial repeated measurement ANOVA with independent factors Block, feature, delay and dependent factor mean number of revisits: F(1, 14) = 19.93, p = 0.001, partial eta-square 0.59) ranging from an average of 1.79 ± 0.17 revisits per switch in participant 4 to 23.40 ± 1.91 revisits per switch for participant 13. For a detailed visualization see chapter VIII. Supplementary: Figure 71.

Now looking at the chunk size and number of fixations separately, it is to note that the chunk size for Block 1 with complex objects and a long delay as well as Block 1 with easy objects and a short delay were not distributed normally (Kolmogorov-Smirnov: $D(15) = 0.22$, $p = 0.041$; $D(15) = 0.23$, $p = 0.033$).

We can see that the chunk size increased in Block 2 (Multifactorial repeated measurement ANOVA with independent factors Block, feature, delay and dependent factor mean chunk size: $F(1, 14) = 32.99$, $p < 0.001$, partial eta-square 0.70), with a mean chunk size of 3.97 ± 0.47 in Block 1 and 5.73 ± 0.69 in Block 2 (Figure 44: A). For every object that was remembered simultaneously in Block 1 1.43 objects were remembered simultaneously in Block 2 (Pearson coefficient: $D(15) = 0.93$, $p < 0.001$, Linear fit: $y = 1.43x$, $R^2 = 0.98$; Figure 44: B).

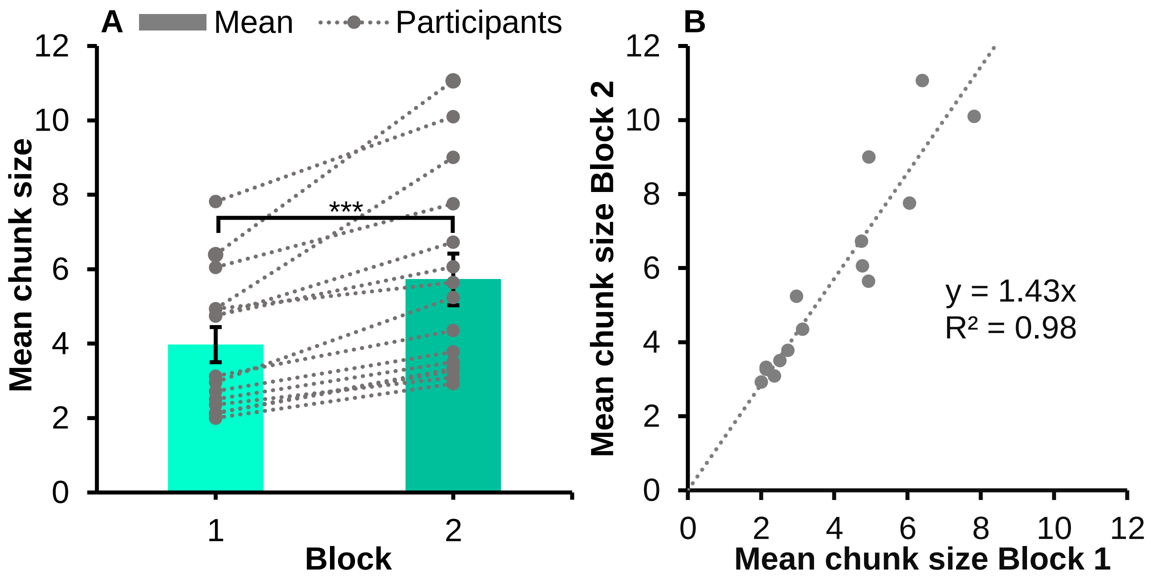

Figure 44: Mean chunk size per Block
A: Mean chunk size plotted against the Block: Block 1 (light) and Block 2 (dark) averaged over participants (n = 15) as bars and for each participant as dots.
B: Mean chunk size in Block 2 plotted against mean chunk size in Block 1 for n = 15 participants.

Neither delay (Multifactorial repeated measurement ANOVA with independent factors Block, feature, delay and dependent factor mean chunk size: $F(1, 14) = 0.14$, $p = 0.71$, partial eta-square 0.010), nor the complexity (Multifactorial repeated measurement ANOVA with independent factors Block, feature, delay and dependent factor mean chunk size: $F(1, 14) = 1.63$, $p = 0.22$, partial eta-square 0.10) influenced the chunk size. Furthermore, the average chunk size was nearly the same for both hemifields (Paired t-test two-sided: $t(14) = 1.28$, $p = 0.22$).

Looking at the number of revisits, it is to note that only the number of revisits for Block 2 in the easy condition with a short delay was distributed normally. (Block 1, complex, long: Kolmogorov-Smirnov: D(15) = 0.240, p = 0.020; Block 1, complex, short: Kolmogorov-Smirnov: D(15) = 0.23, p = 0.029; Block 1, easy, long: Kolmogorov-Smirnov: D(15) = 0.240, p = 0.020; Block 1, easy, short: Kolmogorov-Smirnov: D(15) = 0.22, p = 0.047; Block 2, complex, long: Kolmogorov-Smirnov: D(15) = 0.27, p = 0.004; Block 2, complex, short: Kolmogorov-Smirnov: D(15) = 0.27, p = 0.004; Block 2, easy, long: Kolmogorov-Smirnov: D(15) = 0.23, p = 0.037).

In Block 2 participants made more revisits compared to Block 1 (Multifactorial repeated measurement ANOVA with independent factors Block, feature, delay and dependent factor mean number of revisits: F(1, 14) = 14.90, p = 0.002, partial eta-square 0.516). In Block 1 the mean number of revisits was 4.38 ± 1.11 was more than doubling in Block 2 with an average of 10.84 ± 2.45 (Figure 45: A). For every revisit in Block 1 more than two revisits were made in Block 2 (Pearson coefficient: D(15) = 0.81, p < 0.001, Linear fit: y= 2.15x, R^2 = 0.84; Figure 45: B). It is to note that the number of revisits was not distributed normally for Block 1 or 2 (Kolmogorov-Smirnov: D(15) = 0.24, p = 0.018; D(15) = 0.22, p = 0.049).

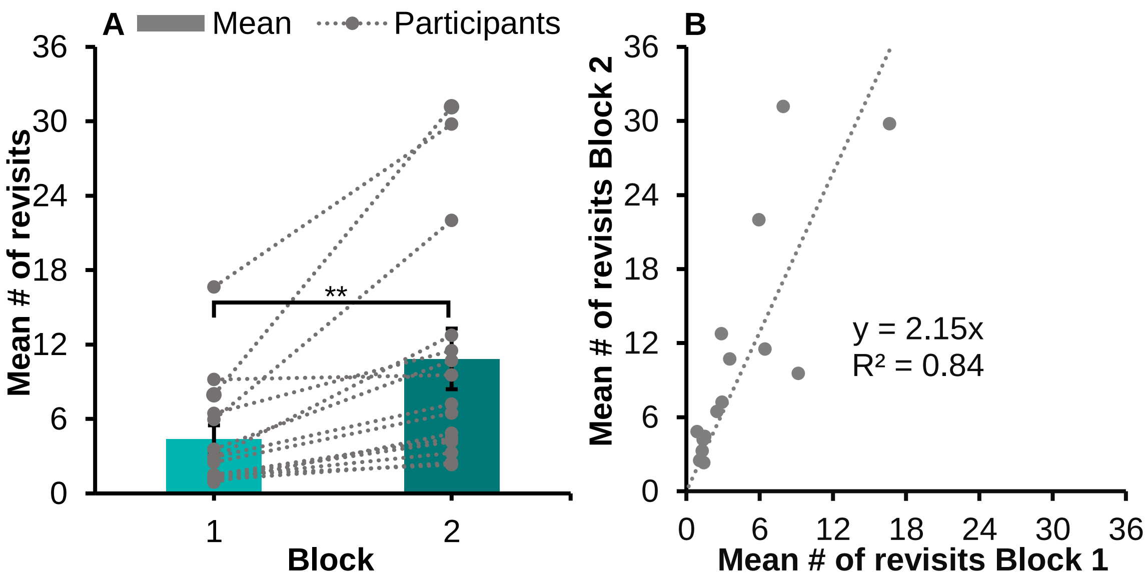

Figure 45: Mean number of revisits per Block
A: Mean number of revisits plotted against the Block: Block 1 (light) and Block 2 (dark) averaged over participants (n = 15) as bars and for each participant as dots.
B: Mean number of revisits in Block 2 plotted against the mean number of revisits in Block 1 for n = 15 participants.

Neither the complexity (Multifactorial repeated measurement ANOVA with independent factors Block, feature, delay and dependent factor mean number of revisits: F(1, 14) = 3.51, p = 0.082, partial eta-square 0.201) nor the delay (Multifactorial repeated measurement ANOVA with independent factors Block, feature, delay and dependent factor mean number of revisits: F(1, 14) = 0.77, p = 0.40, partial eta-square 0.052) did affect the number of revisits.

However, there was an interaction between the Block and the complexity regarding the number of revisits (Multifactorial repeated measurement ANOVA with independent factors Block, feature, delay and dependent factor mean number of revisits: F(1, 14) = 5.70, p = 0.032, partial eta-square 0.29). In Block 2 participants needed more revisits for complex objects compared to simple ones (Figure 46).

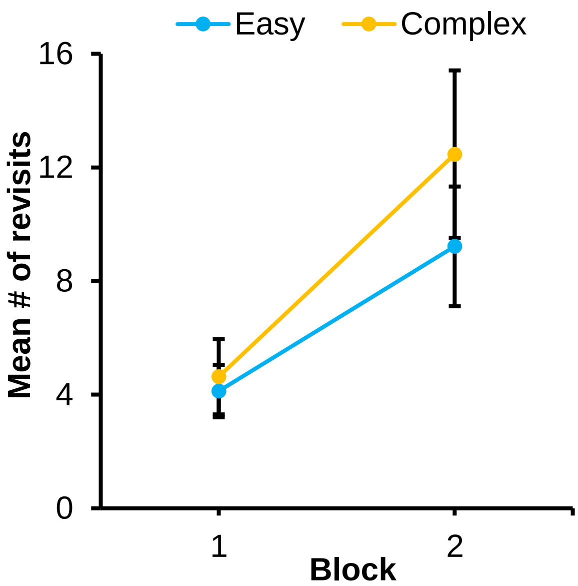

Figure 46: Interaction of Block and complexity on mean number of revisits
Mean number of revisits is plotted against the Block for both conditions: easy (blue) and complex (yellow), averaged over all participants (n = 15).

Different from the chunk size, the number of revisits was higher on the left side with 7.31 ± 1.66 compared to the right side with 6.02 ± 1.41 revisits on average (Paired two-sided t-test: t (14) = 3.01, p = 0.009; Figure 47: A). For every revisit done on the left side only 0.83 revisits were done on the right side (Pearson coefficient: D(15) = 0.97, p < 0.001, Linear fit: y= 0.83, R^2 = 0.98; Figure 47: B).

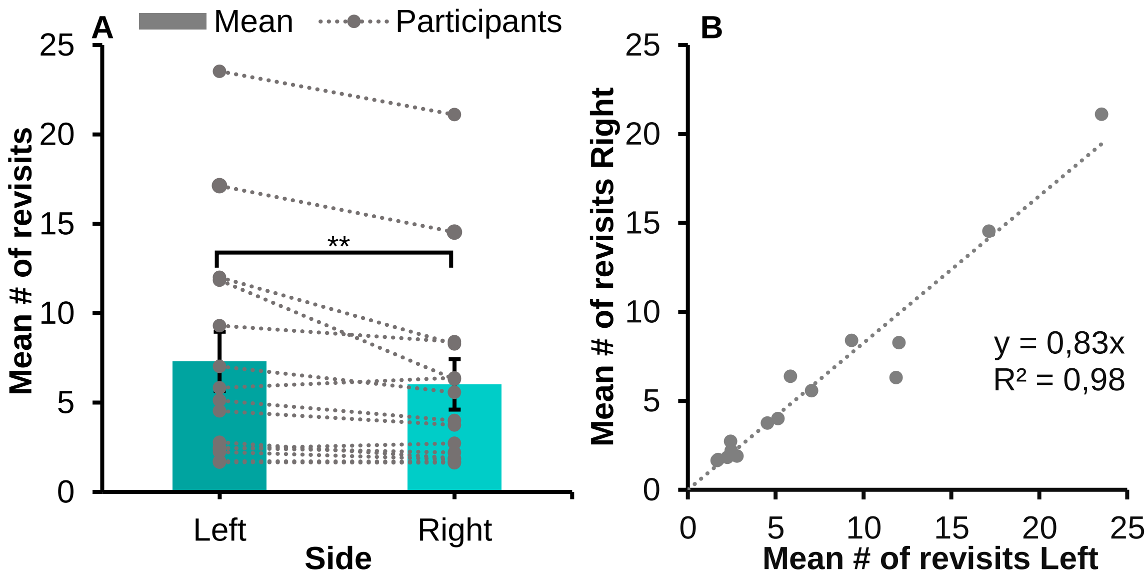

Figure 47: Mean number of revisits depending on screen side
A: Mean number of revisits plotted against the screen side: right (light blue) and left (dark blue) averaged over all participants (n = 15) as bars as well as for each participant as dots.
B: Correlation between the mean number of revisits between the left and the right side plotted for all participants (n = 15).

Looking at the parameter chunk size and number of revisits together, we already saw a correlation between both of them for all participants (Figure 43). To assess the nature of this correlation some principles must be set. First, we assume that the chunk size affects the number of revisits. Second the chunk size cannot be smaller than 1 or else the task could not be solved. Furthermore, due to the constraints of the working memory, chunk size must be limited. Because participants had to solve the task as fast and as good as possible, the time they had solving the task and therefore also the number of revisits may have been restricted too.

The correlation between chunk size and number of revisits can also be examined pooled over all switches and participants and not just on the level of participants as done before. On a descriptive level one can see that the number of revisits increased with chunk size (Figure 48). Furthermore, the number of revisits per chunk size was not distributed normally for most chunk sizes. Only the number of revisits for chunk size 15 was normally distributed (for details see chapter VIII. Supplementary: Table 4).

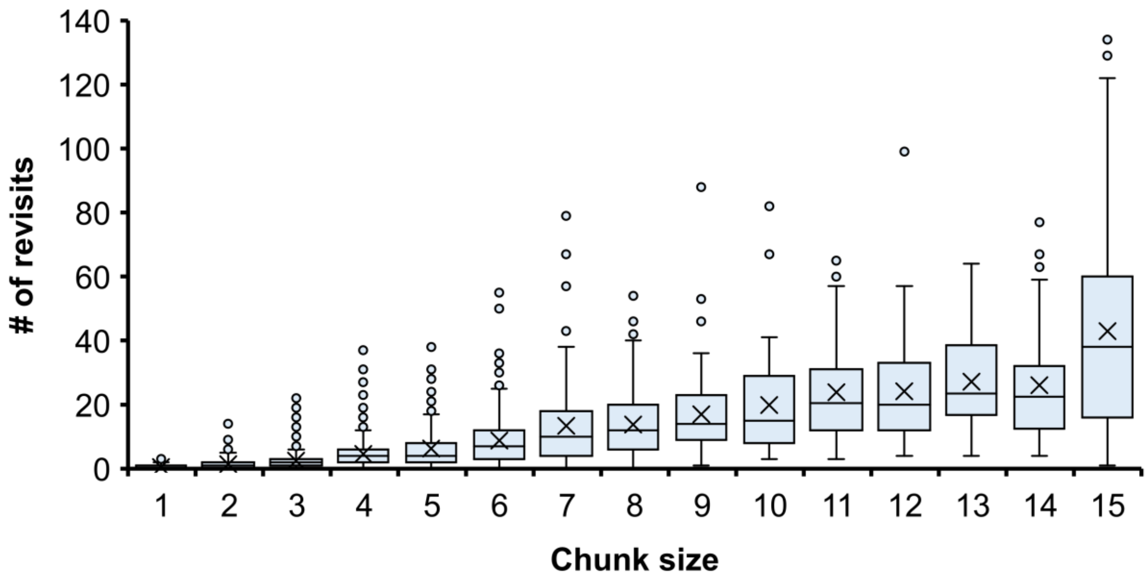

Figure 48: Number of revisits for all chunk sizes
Boxplot of the number of revisits plotted for each chunk size (n = 15) pooled over all participants and switches. Circles symbolize statistical outliers, crosses mark the mean, while the median is visualized as a line.

In order to assess if the median or a corrected mean would be a better parameter to use, all three measurements are compared in the following. Using the median elicited a problem as the median number of revisits was zero for a chunk size of one ruling a power fit impossible. Nevertheless, the median number of revisits increased with chunk size (Pearson correlation coefficient: $r = 0.96$, $p < 0.001$, $n = 15$). There was nearly no difference between the normal mean (Pearson correlation coefficient: $r = 0.97$, $p < 0.001$, $n = 15$, Power fit: $y = 0.47\ x^{1.61}$, $R^2 = 0.99$) and a corrected mean (Pearson correlation coefficient: $r = 0.96$, $p < 0.001$, $n = 15$, Power fit: $y = 0.37\ x^{1.68}$, $R^2 = 0.99$), in which statistical outliers were excluded (Figure 49). When excluding outliers, the small differences between the normal mean and the corrected one become more insignificant. By including the previously shown high variability in the number of revisits at higher chunk sizes, the decision to use the normal mean seems feasible even though the data is not distributed normally.

Because the difference between corrected and normal mean is rather small, the normal mean will be used in the following. As seen before the mean number of revisits per chunk size differed between the four combinations: easy short, easy long, complex short and complex long and increased with chunk size. For conditions with complex objects the increase in number of revisits was steeper (Complex Short Pearson correlation coefficient: r = 0.94, p < 0.001, n = 15, Power fit: $y = 0.41 \times ^{1.73}$, $R^2 = 0.97$; complex long Pearson correlation coefficient: r = 0.97, p < 0.001, n = 15, Power fit: $y = 0.65 \times ^{151}$, $R^2 = 0.99$) compared to easy objects (Easy Short Pearson correlation coefficient: r = 0.91, p < 0.001, n = 15, Power fit: $y = 0.37 \times ^{1.68}$, $R^2 = 0.98$; Complex Long Pearson correlation coefficient: r = 0.88, p < 0.001, n = 15, Power fit: $y = 0.41 \times ^{1.61}$, $R^2 = 0.98$). Similar tendencies could be seen for the delay: in trials with a shorter delay the number of revisits increased steeper compared to longer delay conditions (Figure 50).

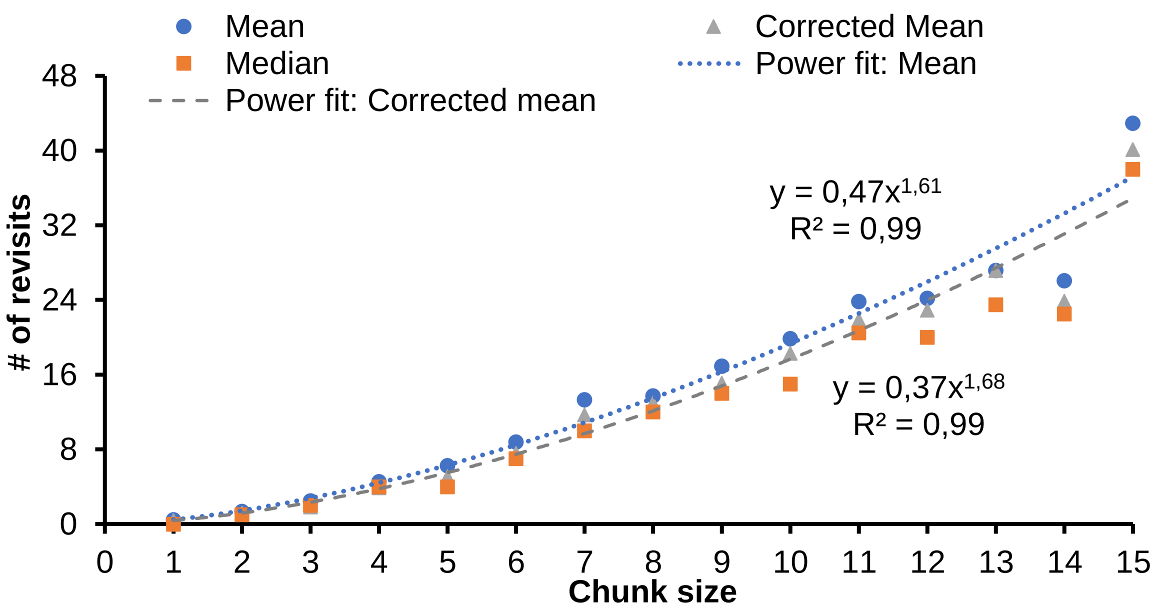

Figure 49: Comparison between mean, corrected mean and median
The mean (blue circles), corrected mean (grey triangles) and median (yellow squares) number of revisits is plotted against the chunk size (n = 15) pooled over all participants and switches. Power fits of the mean (dotted blue line) and the corrected mean (dashed grey line) are included.

By using the mean, it is not possible to make any claims about the abundance of the observations for each chunk size. Therefore, we will now look at the number of observations per chunk size as well as the four different complexity and delay combinations. The most common chunk size used by participants was 2 and the number of observations per chunk size so the incidence of a certain chunk size decreased from there on (Figure 51). For complex objects the number of observations per chunk size was larger with 572 observations for a chunk size of two in the complex long condition and 536 in the complex and short condition, compared to 377 and 365 observations for a chunk size of two in the easy long and easy short condition.

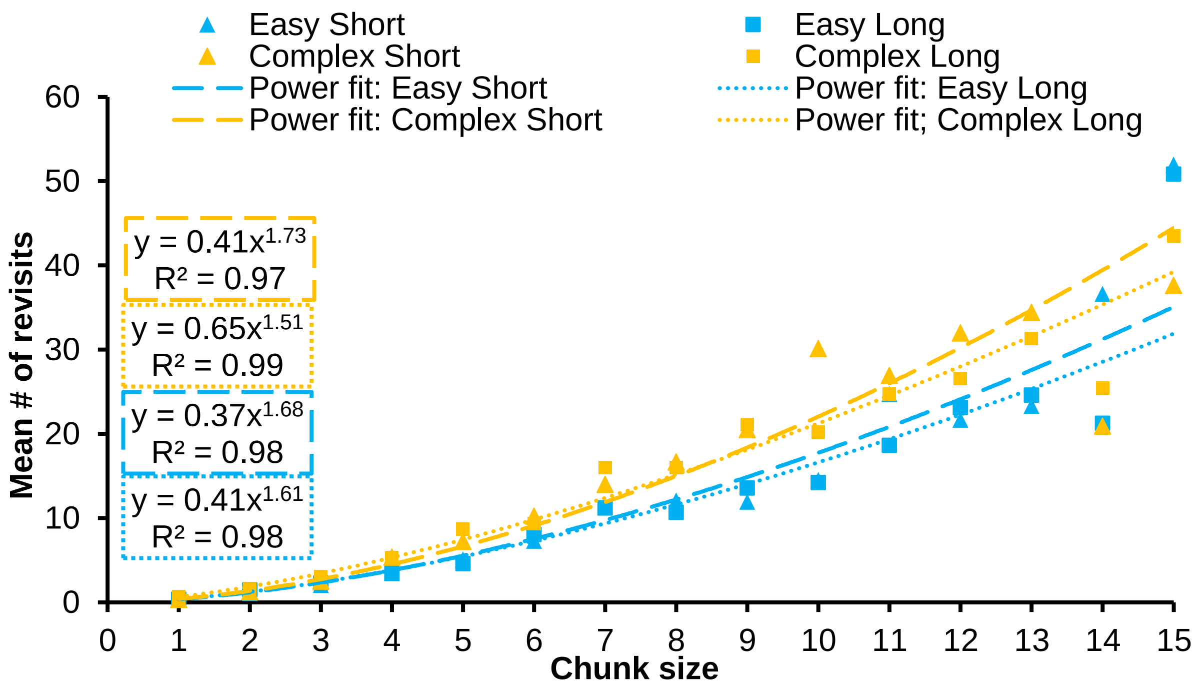

Figure 50: Mean number of revisits for the different chunk sizes
Mean number of revisits plotted per chunk size plotted against the chunk size(n = 15) pooled over all participants and switches or four conditions: Easy short (blue triangle and continuous line), easy long (blue square and dotted line), complex short (yellow triangle and continuous line) and complex long (yellow square and dotted line). Power fits in respective colors are included.

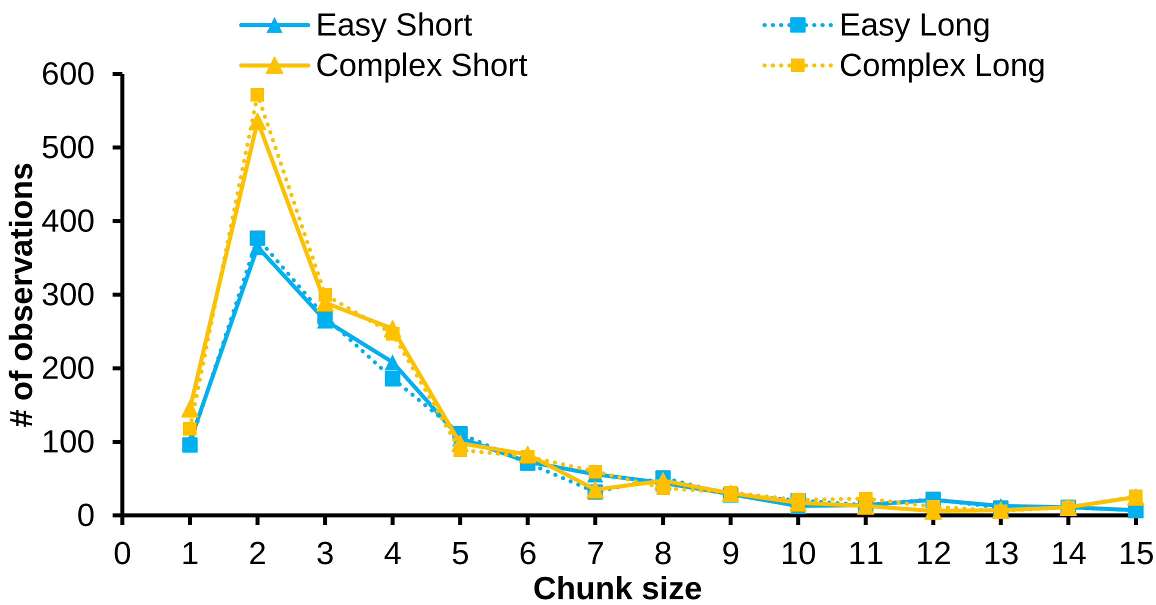

Figure 51: Number of observations per chunk size
The number of observations is plotted against the chunk size (n = 15) pooled over all participants and switches or four conditions: Easy short (blue triangle and continuous line), easy long (blue square and dotted line), complex short (yellow triangle and continuous line) and complex long (yellow square and dotted line).

Looking solely at the number of revisits per chunk size we can see that there were more refixations to complex objects compared to simple objects. While the number of observations was the largest for a chunk size of two, the number of revisits was the largest in switches with a chunk size of four. Furthermore, in trials with complex objects there were many revisits in switches with a chunk size of 15 (Figure 52). In the complex long condition, there were a total of 1131 revisits in switches with a chunk size of 15 and 940 in the complex short condition.

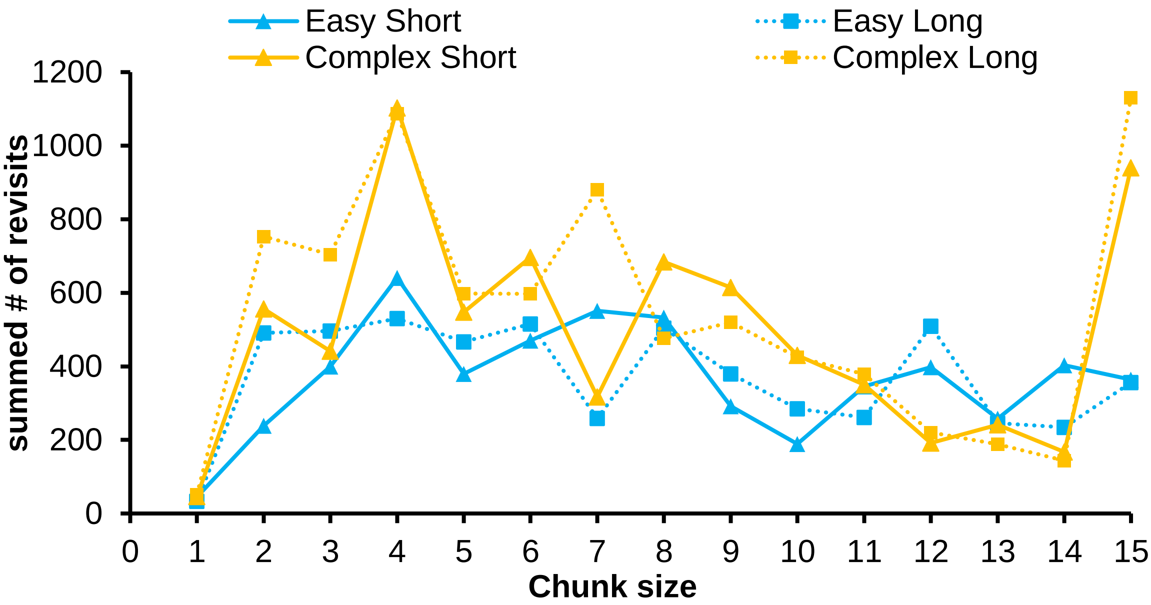

Figure 52: Summed number of revisits per chunk size
The summed number of revisits is plotted against the chunk size (n = 15) pooled over all participants and switches or four conditions: Easy short (blue triangle and continuous line), easy long (blue square and dotted line), complex short (yellow triangle and continuous line) and complex long (yellow square and dotted line).

3.5.2 Chunk and revisit fixation duration

Using the classification of fixations into chunk and revisit based on a switch we can assess whether those fixations differ in their mean fixation duration. It is to note that the fixation duration was not distributed normally in three conditions (Block 2, chunk fixation, complex features and a short delay: Kolmogorov-Smirnov: $D(15) = 0.25$, $p = 0.013$; Block 2, revisit fixations, complex and long delay: Kolmogorov-Smirnov: $D(15) = 0.28$, $p = 0.003$; Block 2, revisit fixations, complex and short delay: Kolmogorov-Smirnov: $D(15) = 0.29$, $p = 0.002$).

As seen before the mean fixation duration differed between the participants (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean fixation duration: $F(1, 14) = 158.68$, $p < 0.001$, partial eta-square 0.92).

The mean fixation duration classified as chunk was on average with 0.50 ± 0.041 s longer compared to a fixation classified as a revisit with 0.44 ± 0.034 s (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean fixation duration: $F(1, 14) = 21.47$, $p < 0.001$, partial eta-square 0.61; Paired two-sided t-test: $t(14) = 4.66$, $p < 0.001$).

It is to note that the mean revisit fixation duration was not normally distributed (Kolmogorov-Smirnov: D(15) = 0.23, p = 0.035). The mean chunk fixation duration was the longest in participant 9 with 0.93 s. Similar tendencies could be seen in the mean revisit fixation duration with 0.83 s for participant 8. The shortest mean chunk fixation duration was 0.32 s for participant 10. Participant 10 also had the shortest mean revisit fixation duration with 0.28 s (Figure 53: A).

For each second a fixation classified as chunk lasted, the fixation duration classified as a revisit lasted only 0.88s (Pearson coefficient correlation coefficient: $r = 0.94$, $p < 0.001$, $n = 15$, Linear fit: $y = 0.85x$, $R^2 = 0.99$; Figure 53: B).

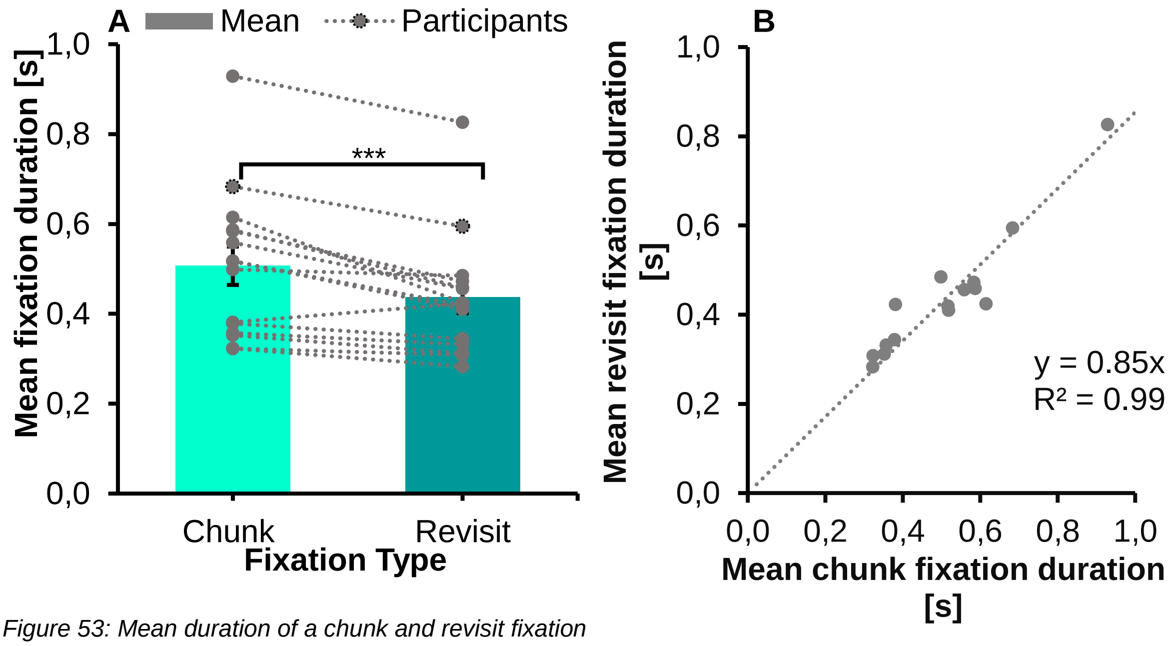

Figure 53: Mean duration of a chunk and revisit fixation
A: Mean fixation duration [s] plotted against the fixation type: chunk (turquoise) or revisit (petrol) averaged over all participants (n = 15) as bars as well as for each participant as dots averaged over all fixations.
B: Mean duration of a revisit fixation [s] plotted against the mean duration of a chunk fixation [s] plotted for all participants (n = 15).

As seen before the Block (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean fixation duration: $F(1, 14) = 11.17$, $p = 0.005$, partial eta-square 0.44; Figure 16: A) and delay (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean fixation duration: $F(1, 14) = 37.36$, $p < 0.001$, partial eta-square 0.73; Figure 17: C) effected the fixation duration. However, by differentiating between a chunk and revisit fixation the previously seen tendency that complexity influenced the fixation duration became significant (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean fixation duration: $F(1, 14) = 4.77$, $p = 0.047$, partial eta-square 0.25; Figure 17: A). The previous seen interaction between Block and delay could be confirmed again (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean fixation duration: F(1, 14) = 8.24, p = 0.012, partial eta-square 0.37; Figure 18: A)

Using the fixation type as an additional factor we can see an interaction between the fixation type and the delay duration (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean fixation duration: F(1, 14) = 43.25, p < 0.001, partial eta-square 0.76). In trials with a long delay the duration of a chunk fixation was longer compared to the duration of a revisit fixation, which stayed approximately the same (Figure 54).

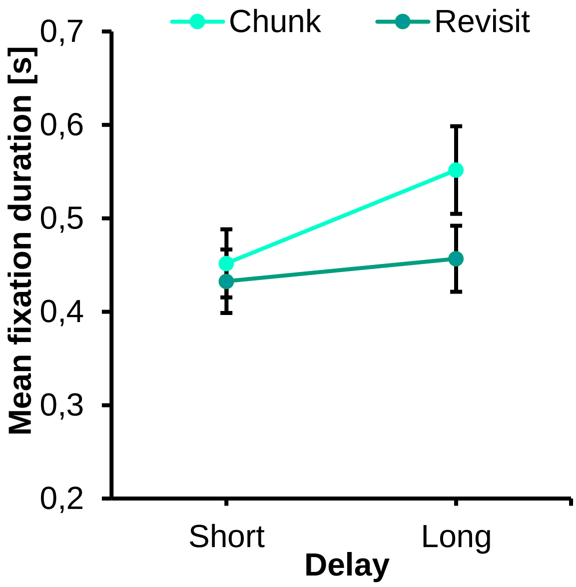

Figure 54: Interaction between delay and fixation type for mean fixation duration
Mean fixation duration [s] is plotted against the delay duration: short and long for both fixations types: chunk (turquoise) and revisit (petrol), averaged over participants (n = 15).

There is no difference in the duration of a chunk or revisit fixation between the left or the right side (Multifactorial repeated measurement ANOVA with independent factors fixation type, side and dependent factor mean fixation duration: F(1, 14) = 0.40, p = 0.54, partial eta-square 0.28).

3.5.3 Number of chunk and revisit fixations per trial

As a reminder how fixations were categorized into chunk or revisit this classification took place on the level of a switch. Each fixation within a switch to a not already fixated object is classified as a chunk fixation, while each refixation of an already fixated object within a switch is classified as a revisit fixation. Even though the classification itself took place on the level of a switch, the subdivision of fixation types is also usable on trial level, which will be examined in the following.

It is to note that six out of the 16 data sets were not distributed normally (Block 1 Chunk complex Short Kolmogorov-Smirnov: D(15) = 0.30, p < 0.001; Block 1 Revisit complex long Kolmogorov-Smirnov: D(15) = 0.27, p = 0.005; Block 1 Revisit easy long Kolmogorov-Smirnov: D(15) = 0.24, p = 0.016; Block 2 Chunk Complex long Kolmogorov-Smirnov: D(15) = 0.24, p = 0.017; Block 2 Chunk Complex Short Kolmogorov-Smirnov: D(15) = 0.27, p = 0.004; Block 2 Revisit Complex long Kolmogorov-Smirnov: D(15) = 0.27, p = 0.004; Block 2 Revisit complex Short Kolmogorov-Smirnov: D(15) = 0.25, p = 0.012).

Between all 15 participants there were on average 38.40 ± 2.30 chunk fixations per trial and 13.40 more revisit fixations with an average of 51.80 ± 8.31 (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, complexity and delay and dependent factor mean number of fixations per trial: $F(1, 14) = 3.72$, $p = 0.074$, partial eta-square 0.21). The mean number of chunk fixations per trial ranged from 28.56 ± 1.24 in participant 5 to 62.13 ± 3.78 in participant 11 (Figure 55: A). Overall participants differed vastly in their number of fixations per trial (Multifactorial repeated measurement ANOVA with independent factors fixation type, side and dependent factor mean fixation duration: $F(1, 14) = 80.43$, $p < 0.001$, partial eta-square 0.85). The mean number of revisits per trial on the other hand ranged from 18.03 ± 0.95 in participant 8 to 110.43 ± 7.69 in participant 13. Both, the mean number of chunk fixations as well as the mean number of revisit fixations per trial are not distributed normally (Kolmogorov-Smirnov: $D(15) = 0.24$, $p = 0.020$; Kolmogorov-Smirnov: $D(15) = 0.28$, $p = 0.002$). With an increase in the number of chunk fixations the number of revisit fixations increased too. However, the nature of this increase is inconclusive (Pearson coefficient correlation coefficient: $r = 0.69$, $p = 0.004$, $n = 15$). While most participants had similar mean numbers of chunk and revisit fixations, 5 participants made more fixations in general especially revisit fixations (Figure 55: B).

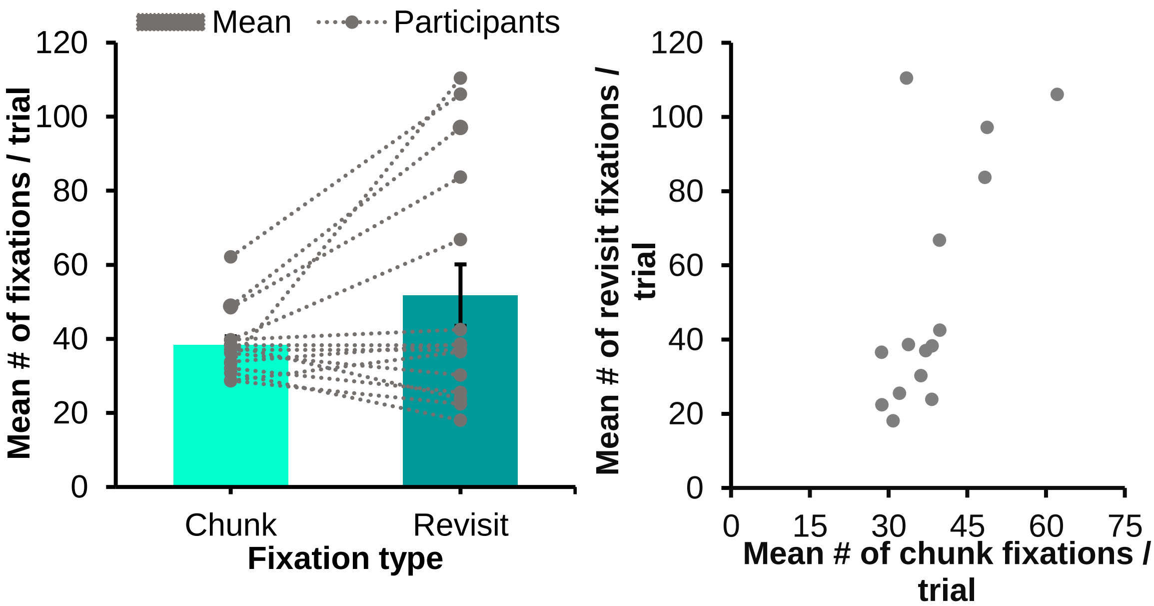

Figure 55: Mean number of chunk and revisit fixations per trial
A: Mean number of fixations per trial is plotted against the fixation type: chunk (turquoise) or revisit (petrol) averaged over all participants (n = 15) as bars as well as for each participant averaged over all trials (n = 40) as dots.
B: Mean number of revisit fixations per trial is plotted against the mean number of chunk fixations per trial for all participants (n = 15) averaged over all trials (n = 40).

As seen before the mean number of fixations per trial differed between the Blocks (Multifactorial repeated measurement ANOVA with independent factors fixation type, side and dependent factor mean fixation duration: $F(1, 14) = 11.34$, $p = 0.005$, partial eta-square 0.448) and different complexities (Multifactorial repeated measurement ANOVA with independent factors fixation type, side and dependent factor mean fixation duration: F(1, 14) = 20.01, p = 0.001, partial eta-square 0.59). The mean number of fixations did not differ for a short or a long delay even when fixations were differentiated between chunk and revisit (Multifactorial repeated measurement ANOVA with independent factors fixation type, side and dependent factor mean fixation duration: F(1, 14) = 2.58, p = 0.13, partial eta-square 0.156).

In addition to the previously found interactions between Block and complexity (Multifactorial repeated measurement ANOVA with independent factors fixation type, side and dependent factor mean fixation duration: F(1, 14) = 6.97, p = 0.019, partial eta-square 0.33) and delay and feature complexity (Multifactorial repeated measurement ANOVA with independent factors fixation type, side and dependent factor mean fixation duration: F(1, 14) = 4.65, p = 0.49, partial eta-square 0.25) new interactions could be found.

First of all, there was an interaction between Block and fixation type (Multifactorial repeated measurement ANOVA with independent factors fixation type, side and dependent factor mean fixation duration: F(1, 14) = 26.21, p < 0.001, partial eta-square 0.65). The number of chunk fixations per trial decreased slightly from Block 1 to Block 2. The number of revisit fixations per trial on the other hand increased by 67 % from 38.74 to 64.8 (Figure 56: A).

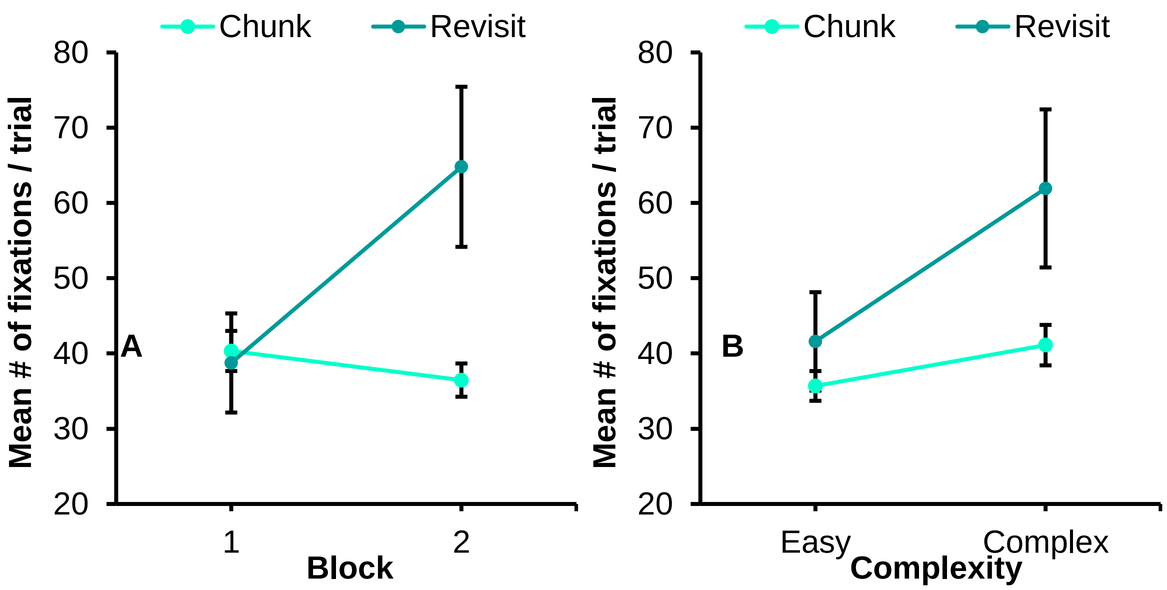

Figure 56: Interactions between fixation type and Block or complexity
A: Mean number of fixations per trial plotted against the Block for both fixation types: chunk (turquoise) or revisit (petrol) averaged over all participants (n = 15) and conditions.
B: Mean number of fixations per trial plotted against the complexity for both fixation types: chunk (turquoise) or revisit (petrol) averaged over all participants (n = 15) and conditions.

Furthermore there was an interaction between type and complexity in the mean number of fixations per trial (Multifactorial repeated measurement ANOVA with independent factors fixation type, side and dependent factor mean fixation duration: F(1, 14) = 8.23, p = 0.012, partial eta-square 0.37). While the mean number of chunk fixations increased slightly in complex trials, the number of revisits per trial had a substantial increase from 41.61 to 61.93 in complex trials (Figure 56: B).

There were two complex interactions. The first interaction was between Block, fixation type and complexity (Multifactorial repeated measurement ANOVA with independent factors fixation type, side and dependent factor mean fixation duration: $F(1, 14) = 18.79$, $p = 0.001$, partial eta-square 0.57). While the mean number of chunk fixations increased similarly between complexities and Blocks, the mean number of revisit fixations per trial had a considerable rise in the complex condition in Block 2 (Figure 57: A).

The second interaction was between Block, fixation type and delay (Multifactorial repeated measurement ANOVA with independent factors fixation type, side and dependent factor mean fixation duration: $F(1, 14) = 7.25$, $p = 0.018$, partial eta-square 0.34). While the number of fixations stayed approximately the same between the two delay conditions for chunk fixations independent of Block number as well as revisit fixations in Block 2, there was an increase in mean number of revisit fixations for Block 1 between the short and the long delay condition (Figure 57: B).

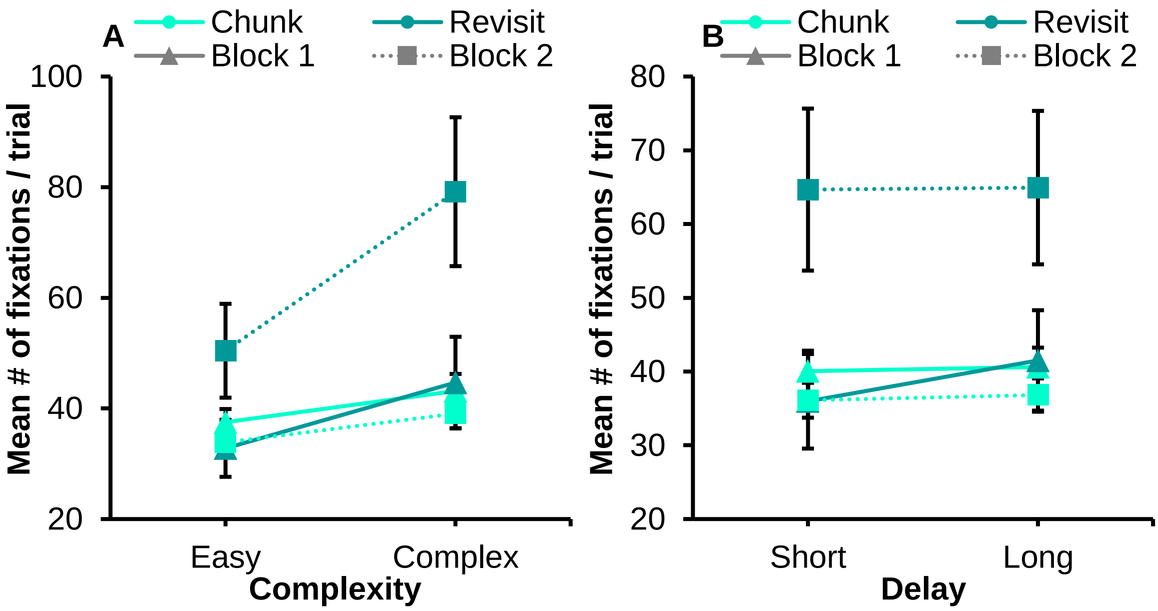

Figure 57: Interaction between Block, fixation type and complexity or delay of the mean number of fixations per trial
A: Mean number of fixations per trial plotted against the complexity for both fixation types chunk (turquoise) or revisit (petrol) and Block: Block 1 as triangles with continuous lines and Block 2 with squares and dotted lines averaged over all participants (n = 15) and respective conditions.
B: Mean number of fixations per trial plotted against the delay for both fixation types chunk (turquoise) or revisit (petrol) and Block: Block 1 as triangles with continuous lines and Block 2 with squares and dotted lines averaged over all participants (n = 15) and respective conditions.

Because we saw a difference between the sides, further analyses were conducted. As seen before, there were more fixations on the left compared to the right side (Multifactorial repeated measurement ANOVA with independent factors fixation side, Block, type, complexity, delay and dependent factor mean number of fixations per trial: $F(1, 14) = 17.38$, $p = 0.001$, partial eta-square 0.55).

In addition to the previously mentioned findings there was an interaction between the side and fixation type (Multifactorial repeated measurement ANOVA with independent factors fixation side, Block, type, complexity, delay and dependent factor mean number of fixations per trial: $F(1, 14) = 13.36$, $p = 0.003$, partial eta-square 0.49). The mean number of revisits on the left side was nearly 30 per trial, while it was only around 22 on the right side. The mean number of chunk fixations on the other hand stayed approximately the same independent of screen side (Figure 58: A).

There was a threefold interaction between side, fixation type and complexity (Multifactorial repeated measurement ANOVA with independent factors fixation side, Block, type, complexity, delay and dependent factor mean number of fixations per trial: $F(1, 14) = 6.05$, $p = 0.028$, partial eta-square 0.302). Including the complexity in this interaction, showed that participants used more fixations in the complex condition. Furthermore, more revisit fixations were made, in general and especially on the left side (Figure 58: B).

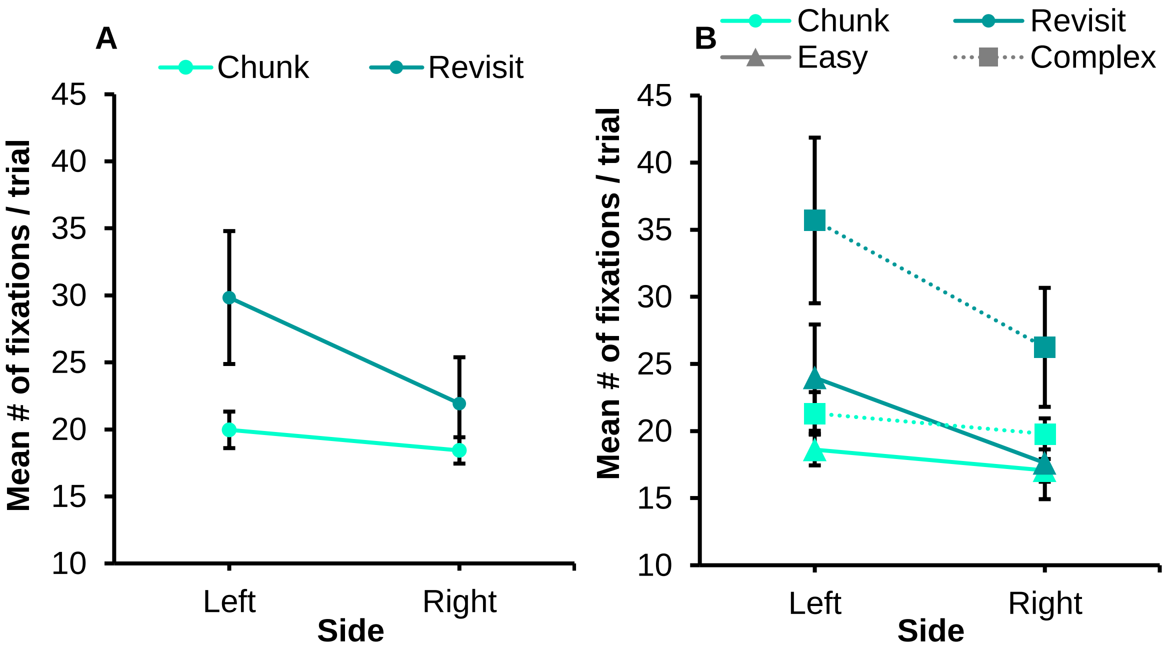

Figure 58: Interaction between Block, fixation type and complexity or delay for the mean number of fixations per trial
A: Mean number of fixations per trial plotted against the side for both fixation types: chunk (turquoise) or revisit (petrol) averaged over all participants (n = 15) and conditions.
B: Mean number of fixations per trial plotted against the side for both fixation types chunk (turquoise) or revisit (petrol) and the complexity: Easy as triangles with continuous lines and complex with squares and dotted lines averaged over all participants (n = 15) and respective conditions.

3.5.4 Total chunk and revisit fixation time per trial

To assess whether the fixation type influenced the mean total fixation time per trial further analysis was conducted. It is to note that four of the used total fixation time combinations are not normally distributed (Block 1, revisit fixation, complex features and a short delay: Kolmogorov-Smirnov: D(15) = 0.26 p = 0.008; Block 1, revisit fixations, complex and long delay: Kolmogorov-Smirnov: D(15) = 0.26, p = 0.007; Block 2, revisit fixation, complex features and a long delay: Kolmogorov-Smirnov: D(15) = 0.24 p = 0.022; Block 2, revisit fixations, easy and short delay: Kolmogorov-Smirnov: D(15) = 0.23, p = 0.030).

As seen before the total fixation time differed between the participants (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean total fixation time: F(1, 14) = 170.06, p < 0.001, partial eta-square 0.92). The mean total fixation time per trial for chunk fixations ranged from an average of 10.91 ± 0.66 s in participant 10 to 28.65 ± 2.38 s for participant 8. Participant 4 had the shortest mean total fixation time of previously visited objects with 9.79 ± 0.82 s per trial, while participant had the longest with an average of 46.70 ± 3.46 s per trial (Figure 59: A).

The mean total fixation time for chunk fixations averaged over all participants was 18.81 ± 1.30 s per trial and for revisit fixations it was 20.54 ± 2.79 s per trial (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean fixation duration: F(1, 14) = 0.30, p = 0.59, partial eta-square 0.021).

The mean revisit total fixation time per trial is not normally distributed (Kolmogorov-Smirnov: D(15) = 0.26 p = 0.007) and not correlated with the average chunk total fixation time per trial (Pearson correlation coefficient: r = -0.052, p = 0.86, n = 15) (Figure 59: B).

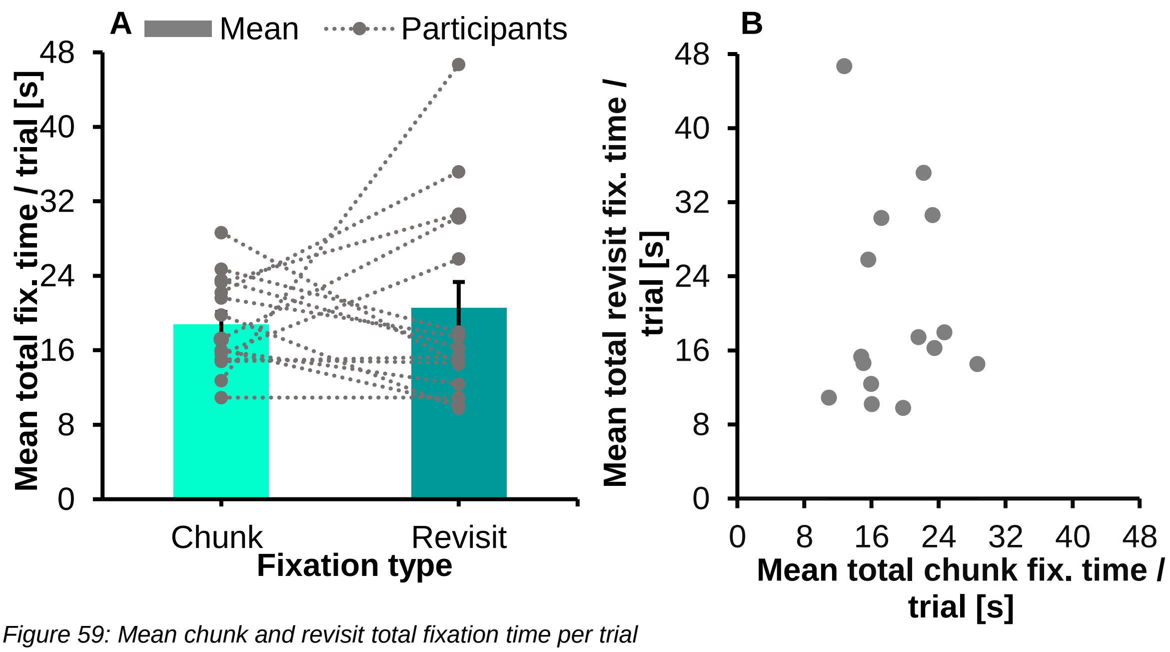

Figure 59: Mean chunk and revisit total fixation time per trial
A: Mean total fixation time [s] plotted against the fixation type: chunk (turquoise) or revisit (petrol) averaged over all participants (n = 15) as bars as well as for each participant averaged over all trials (n = 40) as dots.
B: Mean revisit total fixation time [s] plotted against the mean chunk Total fixation time [s] plotted for all participants (n = 15).

Similar to previous findings the complexity (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean total fixation time per trial: $F(1, 14) = 50.73$, $p < 0.001$, partial eta-square 0.78; see Figure 26: B) as well as the delay (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean total fixation time per trial: $F(1, 14) = 46.57$, $p < 0.001$, partial eta-square 0.77; see Figure 26: A) influenced the mean total fixation time per trial. As seen before there was no difference in mean total fixation time per trial depending on the Block (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean total fixation time per trial: $F(1, 14) = 1.26$, $p = 0.28$, partial eta-square 0.083). This analysis supports previous findings of an interaction between Block and delay (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean fixation duration: $F(1, 14) = 21.61$, $p < 0.001$, partial eta-square 0.61; Figure 27: A) as well as complexity and delay (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean total fixation time per trial: $F(1, 14) = 10.07$, $p = 0.007$, partial eta-square 0.42; Figure 27: B).

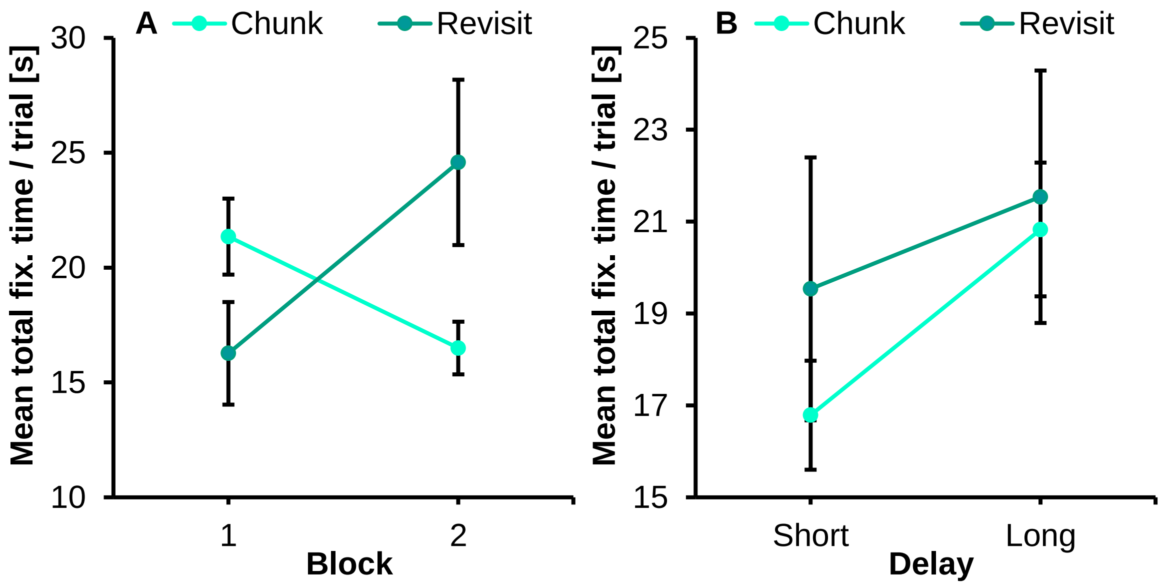

Figure 60: Interactions between fixation type and Block or delay for mean total fixation time per trial
A: Mean total fixation time per trial [s] plotted against the Block for both fixation types: chunk (turquoise) or revisit (petrol) averaged over all participants (n = 15) and conditions.
B: Mean total fixation time per trial [s] plotted against the delay for both fixation types: chunk (turquoise) or revisit (petrol) averaged over all participants (n = 15) and conditions.

Including the fixation type however led to some new findings. We can see that there was an interaction between the Block and the fixation type (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean total fixation time per trial: $F(1, 14) = 34.53$, $p < 0.001$, partial eta-square 0.71). While there was no difference in total fixation time per trial between the Blocks, when the fixation type was not included we can see that the mean chunk total fixation time decreased approximately 5 s from Block 1 to Block 2 while the mean revisit total fixation time increased by nearly 8 s from Block 1 to Block 2 (Figure 60: A).

Furthermore, there was an interaction between the fixation type and the delay (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean total fixation time per trial: $F(1, 14) = 9.37$, $p = 0.008$, partial eta-square 0.40). While the mean revisit total fixation time increased only 2 s between short and long delay trials, the mean chunk total fixation time per trial increased double this time between short and long delays (Figure 60: B).

While there was no interaction between Block and complexity by themselves in the previous and this analysis, there was an interaction between Block and complexity when the fixation type was included in the interaction (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay and dependent factor mean total fixation time per trial: $F(1, 14) = 21.59$, $p < 0.001$, partial eta-square 0.61). The total fixation time per trial increased the most for revisit fixations in trials with complex features in Block 2 (Figure 61). While the increase for the three other combinations was on average 5.83 s the increase of total fixation time per trial for revisit fixations was nearly double this time with 11.96 s.

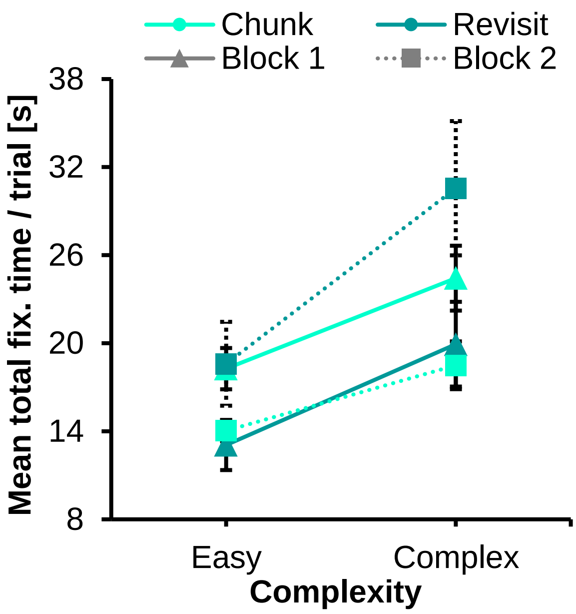

Figure 61: Interaction between Block, complexity and fixation type for the total fixation time per trial. Mean total fixation time per trial [s] plotted against the complexity for both fixation types chunk (turquoise) or revisit (petrol) and Block: Block 1 as triangles with continuous lines and Block 2 with squares and dotted lines averaged over all participants (n = 15) and respective conditions.

Because there was a difference in total fixation time per trial depending on the screen side further analysis was conducted including the fixation type and screen side.

In addition to previous findings there was an interaction between the fixation type and the side (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay, side and dependent factor mean total fixation time per trial: $F(1, 14) = 12.31$, $p = 0.003$, partial eta-square 0.47).The duration of revisit fixations was nearly 3 seconds longer per trial on the left screen side, while the difference for chunk fixations was only 0.59 s between sides (Figure 62: A).

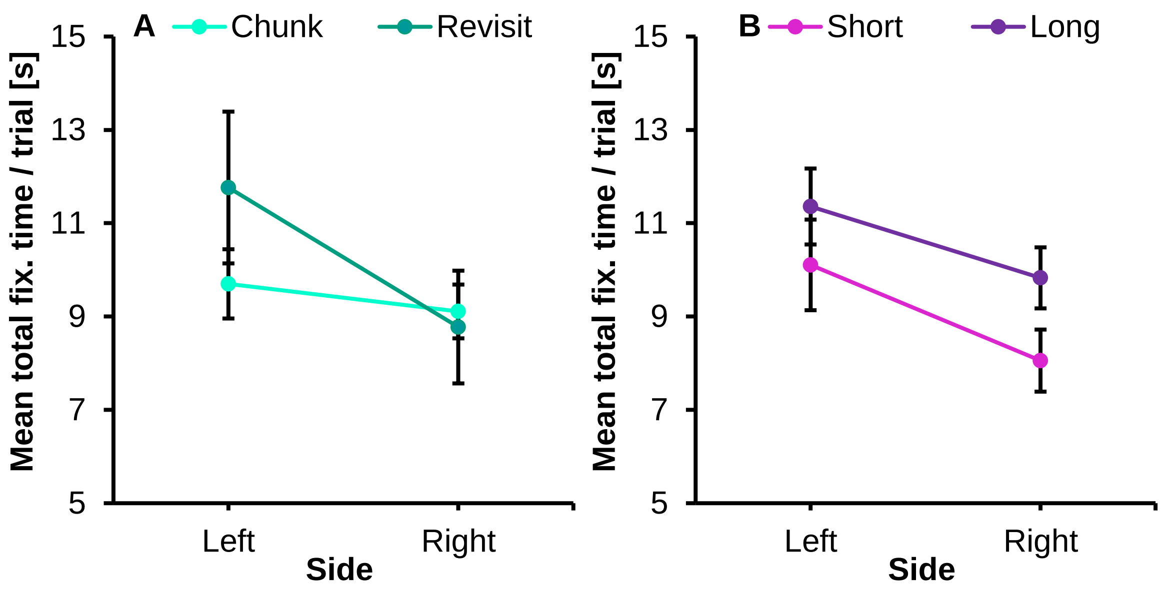

Figure 62: Interaction between side and fixation type or delay for the mean fixation duration per trial
A: Mean total fixation time per trial [s] plotted against the side for both fixation types: Chunk (turquoise) or revisit (petrol) averaged over all participants (n = 15) and conditions.
B: Mean total fixation time per trial [s] plotted against the side for both delay durations: Short (pink) or long (purple) averaged over all participants (n = 15) and conditions.

Furthermore, there was a tendency for an interaction between the delay and the screen side regarding the mean total fixation time per trial (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay, side and dependent factor mean total fixation time per trial: $F(1, 14) = 3.83$, $p = 0.071$, partial eta-square 0.22). While in trial with a long delay the difference in mean total fixation time per trial between the sides was only 1.53 s longer on the left, this difference increased to 2.10 s in trials with a short delay (Figure 62: B).

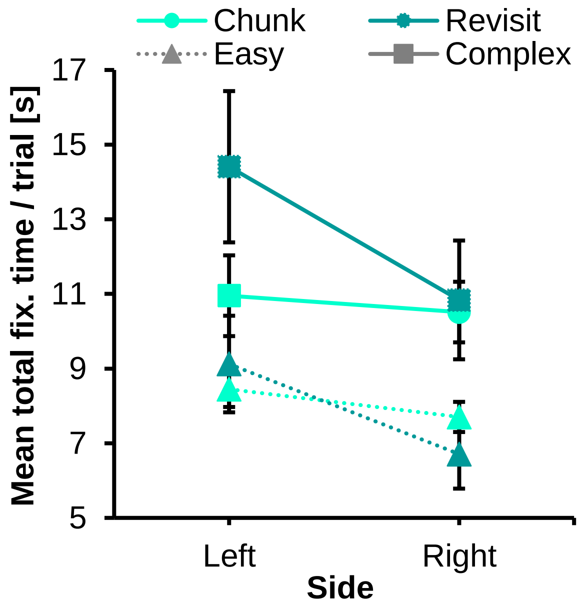

Figure 63: Interaction between screen side, fixation type and object complexity regarding the mean total fixation time per trial
Mean total fixation time per trial [s] plotted against the screen side for both fixation types chunk (turquoise) or revisit (petrol) and complexities: easy as triangles with dotted lines and complex with squares and continuous lines averaged over all participants (n = 15) and respective conditions.

Additionally, there was a rather complex interaction between screen side, complexity and fixation type regarding the mean total fixation time per trial (Multifactorial repeated measurement ANOVA with independent factors fixation type, Block, feature, delay, side and dependent factor mean total fixation time per trial: $F(1, 14) = 5.35$, $p = 0.036$, partial eta-square 0.28). The mean total fixation time per trial was especially long for revisit fixations on the left side in trials with complex objects (Figure

63). Once again one can see that the mean total fixation time per trial was lower in easy trials.

3.6 Correlations

With the main parameters examined the investigation continues by looking at correlations. It is to note that some parameters were bound by definition or logic not included in the fittings but explained in the following.

The first parameter that will be examined is the fixation duration. With an increase in mean fixation duration the mean number of fixations decreased for trials (Pearson correlation coefficient: $r = -0.59$, $p = 0.022$, $n = 15$; Power fit: $y = 41.99x^{-0.88}$, $R^2 = 0.43$; Figure 64: A) as well as switch level (Pearson correlation coefficient: $r = -0.61$, $p = 0.016$, $n = 15$; Power fit: $y = 2.95x^{-1.62}$, $R^2 = 0.53$; Figure 64: B).

The mean fixation duration is limited by the predefined shortest duration a fixation can have. Furthermore, the mean number of chunk fixations per trial is limited to 30 because there were 15 objects per side that had to be compared. Similarly, the mean number of fixations per switch had to be at least one in order to fulfill the task.

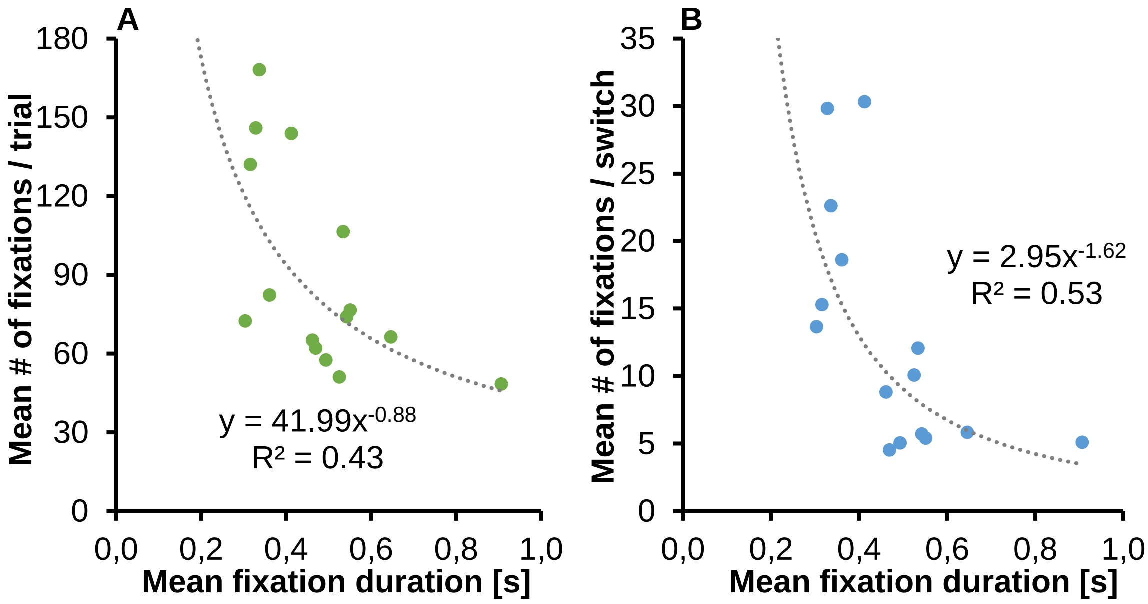

Figure 64: Correlations with mean fixation duration
A: Mean number of fixations per trial is plotted against the mean fixation duration [s] for all participants (n = 15) averaged over all trials (n = 40). The dotted line visualizes the power fit.
A: Mean number of fixations per switch is plotted against the mean fixation duration [s] for all participants (n = 15) averaged over all trials (n = 40). The dotted line visualizes the power fit.

Chunk size decreases with an increase in duration of a chunk fixation (Pearson correlation coefficient: $r = -0.66$, $p = 0.007$, $n = 15$; Power fit: $y = 2.03x^{-1.09}$, $R^2 = 0.53$; Figure 65: A). Furthermore, the duration of a chunk fixation seems to be better suited to predict the number of fixations per switch than the overall (Pearson correlation coefficient: $r = -0.66$, $p = 0.008$, $n = 15$) or revisit fixation duration. However the mean number of revisits also decreased with an increase in duration of such a fixation (Pearson correlation coefficient: r = -0.53, p = 0.041, n = 15; Power fit: $y = 0.97x^{-2.08}$, $R^2 = 0.47$; Figure 65 B).

While the chunk size is limited due to working memory constraints the number of revisits is not limited in such a way. However, one could assume that due to the task description to solve a task as fast and as good as possible it could be possible that participants limited the number of revisits in some way.

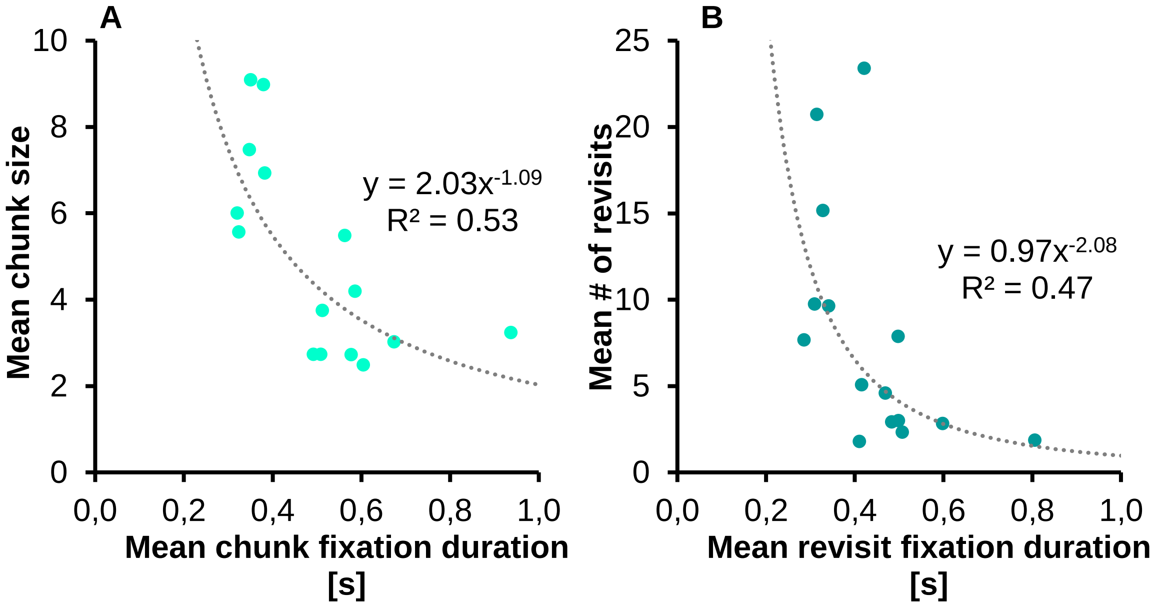

Figure 65: Correlations with mean chunk and revisit fixation duration
A: Mean chunk size is plotted against the mean duration of a chunk fixation [s] for all participants (n = 15) averaged over all trials (n = 40). The dotted line visualizes the power fit.
B: Mean number of revisits is plotted against the mean duration of a revisit fixation [s] for all participants (n = 15) averaged over all trials (n = 40). The dotted line visualizes the power fit.

Looking at the mean number of switches per trial, we can see that the mean number of switches per trial is correlated with the number of fixations per switch (Pearson correlation coefficient: r = -0.68, p = 0.006, n = 15; Power fit: $y = 152.77x^{-1.21}$, $R^2 = 0.55$; Figure 66: A) and the total fixation time per switch (Pearson correlation coefficient: r = -0.68, p = 0.005, n = 15; Power fit: $y = 36.08x^{-0.93}$, $R^2 = 0.55$; Figure 66: B).

Furthermore, chunk size (Pearson correlation coefficient: r = -0.79, p < 0.001, n = 15; Power fit: $y = 35.82^{-0.94}$, $R^2 = 0.70$; Figure 66: C) and number of revisits (Pearson correlation coefficient: r = -0.61, p = 0.017, n = 15; Power fit: $y = 125.19x^{-1.40}$, $R^2 = 0.55$; Figure 66: D) showed similar tendencies. From those four correlated variables chunk size had the biggest effect on the mean number of switches per trial.

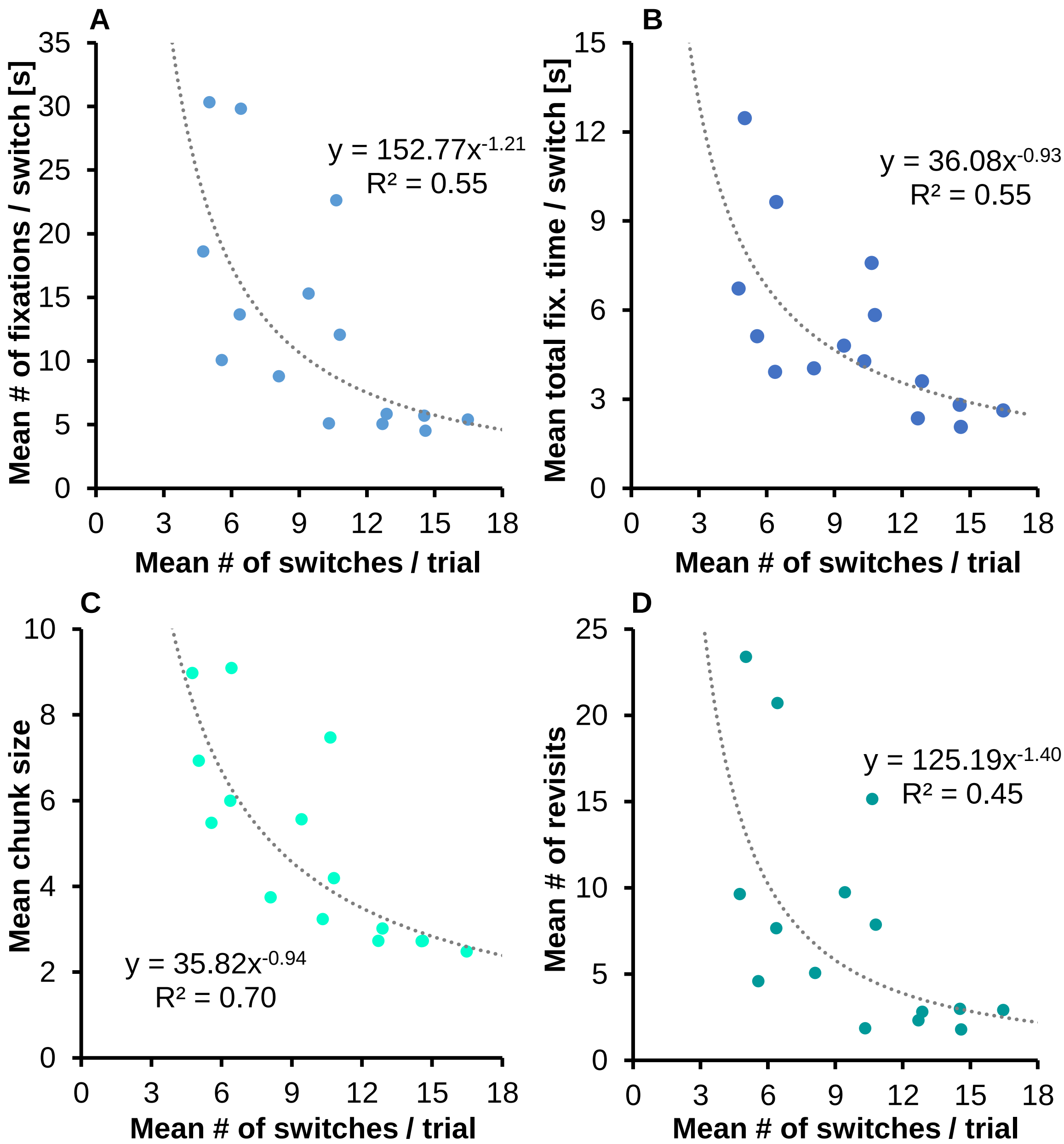

Figure 66: Correlations with mean number of switches per trial
A: Mean number of fixations per switch is plotted against the mean number of switches per trial for all participants (n = 15) averaged over all trials (n = 40). The dotted line visualizes the power fit.
B: Mean Total fixation time per switch [s] is plotted against the mean number of switches per trial for all participants (n = 15) averaged over all trials (n = 40). The dotted line visualizes the power fit.
C: Mean chunk size is plotted against the mean number of switches per trial for all participants (n = 15) averaged over all trials (n = 40). The dotted line visualizes the power fit.
D: Mean number of revisits is plotted against the mean number of switches per trial for all participants (n = 15) averaged over all trials (n = 40). The dotted line visualizes the power fit.

4 Discussion

4.1 Results discussion

4.1.1 Fixation duration

In our experiment the mean fixation duration was around 0.48 ± 0.041 s and did not differ largely between objects or non-objects. This is longer compared to the average of 207.2 ms per fixation found by Pomplun et al. (2001) or fixation durations between 200 ms and 235 ms from Galpin and Underwood (2005). A possible explanation for our finding is that fixations were added when they went to the same object and the gaze shift was small.

The fixation duration changed between the two Blocks with Block 2 having shorter durations corresponding to a decrease to 83% of the fixation duration compared to Block 1. The mean fixation duration was furthermore modulated by the delay duration and by including the differentiation between chunk and revisit also by the object complexity. It is known that the fixation duration can be modulated, corresponding to a strategic adaption based on task demands (Over et al., 2007, Scinto et al., 1986). In a meta-analysis of 11 experiments Moffitt (1980) found that in 7 out of 11 experiments fixation duration depended on the experimental condition. He furthermore stated that in high density displays the dependency of fixation duration on the task was decreased. However, the present experiment had no high-density display and the visual span was predefined by the fixation classifying method to be one, so this effect probably did not come into play.

While the fixation duration took longer in trials with a long delay, it took especially longer in Block 1 compared to Block 2, suggesting that the presence of a switch limitation would increase the difficulty.

Additionally, there was a cross interaction between easy and complex trials regarding the mean fixation duration depending on the Block. While the fixation duration was lower for complex trials in Block 1 it increased by nearly two seconds in Block 2. However, a contrary effect could be seen in easy trials, where the mean fixation duration decreased by nearly one second from Block 1 to Block 2. These findings suggest that participants did not operate at their limit in easy trials but had to switch strategies in order to suffice for the limitation in switches and complex objects being perceptually and mnemonic more challenging.

By including the classification as chunk or revisit, we saw that a chunk fixation was 13.6% longer than a revisit fixation. While the duration of a revisit fixation stayed approximately the same independent of delay duration, the duration of a chunk fixation increased from an average of 0.45 s to an average of 0.55 s in the long delay condition. This increase in mean duration of a chunk fixation of 100 ms indicates a difference in processing of first object fixations within a switch depending on the cognitive demand. Refixations of the objects did not differ, indicating another underlying mechanism. While the first fixation on an object within a switch could be seen as the acquisition of information, refixations may be used to maintain and reinforce the information.

Similar to Gould (1973) it was found that the fixation duration increases with memory load. In case of this study memory load can be estimated by chunk size. Furthermore, complex objects and a long delay increase the demands on memory and had also shown an increase in the fixation duration. These findings are consistent with early findings form Salthouse and Ellis (1980) and Salthouse et al. (1981).

In contrast to literature we found no difference in fixation duration between the objects. However, the ordinality of fixations was not included in the analysis so no statement relating to a coarse-to-fine algorithm could be done. A coarse-to-fine algorithm is based on the assumption that short fixation durations and large saccade amplitudes are used at the start of a trial in order to get an overview of the scene. Later in a trial the fixation duration increases and the saccade amplitudes decreases when a target could not be found in order to increase the needed saliency (Over et al., 2007). This provides the capability to perceive the objects in more detail. Similar effects could also be found outside of a visual search task as proposed by Antes (1974).. However, Over et al. (2007) found that the effects of such a coarse-to-fine algorithm ceases when target conspicuities are constant. In contrast to Scinto et al. (1986) they nevertheless found an increase in fixation duration with the ordinality of the fixation. In our case participants knew where to look for the targets and how the targets will look like so conspicuity was rather constant and no coarse-to-fine algorithm had to be implemented in order to solve the task. Furthermore, the participants were instructed to search for the differences by working down the display.

4.1.2 Number of fixations

Regarding the number of fixations, it could once again be seen that our sorting method was sufficient as there were on average nearly 77 more object fixations to per trial compared to non-object fixations. Participants fixated objects on average 90 times per trial and nearly 12 times per switch. In their experiment Pomplun et al. (2001) found that participants made 2.45 fixations before shifting over to the other hemifield. This big difference could result from the limitations in the number of shifts participants could use in Block2. Similarly, Galpin and Underwood (2005) found around 20 fixations per trial, which is more similar to what we found in a single hemifield switch. However, they had only 10 in contrast to our 10 objects and the objects used by them were all unique, so easier to distinguish.

While the mean fixation duration decreased in Block 2, the mean number of fixations increased by a total of 20 fixations per trial and by nearly 89% per switch. With an average of only 8.35 fixations before a hemifield switch this number seems more similar to the findings from Pomplun et al. (2001).

Furthermore, the mean number of fixations per trial increased from around 77 fixations per trial in easy trials to an average of 103 fixations in complex trials. This corresponds to an increase of 34% regarding the number of fixations from easy to complex trial for every fixation done. Interestingly, the delay duration did not affect the number of fixations per trial. Other parameter than the number of fixations changed with increased delay duration indicating the expected increase in difficulty. Regarding the number of fixations per switch, neither the delay nor the feature complexity had an effect on the number of fixations per switch.

Similar to our findings regarding the mean fixation duration we saw that participants did more fixations especially in trials with complex objects in order to suffice the limitation of switches in Block 2. There was also a slight increase in the number of fixations per trial between the Blocks for easy trials, suggesting a change in the used strategy.

While the number of fixations stayed approximately the same independent of delay duration in trials with easy objects, participants made more fixations in trials with complex objects especially when the delay was long. This suggests a bigger cognitive demand in such a condition, which was counteracted by a behavioral adaptation.

While the fixation duration did not differ between the hemifields, participants fixated objects on the left more often compared to objects on the right side. This finding could resemble a difference in cognitive processes between the hemifields. Because each trial started on the left side, participants had to encode these objects first, memorize them during the delay and then compare them with the objects visible on the other side. Assuming this hypothesis is correct, the difference between the sides would resemble the increased demand in encoding as well as memorization compared to decoding and comparison. The finding that the difference in sides results mainly from a difference in the number of fixations between the first four objects contradicts this hypothesis. It is possible that this finding resembles an effect found by Over et al. (2007) in which participants tend to fixate the first object longer than the following. In our case this effect relates to the number of fixations not the fixation duration itself and extends to all objects within the first switch.

By including the classification of fixations as revisit and chunk some new insights could be acquired. Even though this difference was not statistically significant participants needed on average 38 chunk fixations and 52 revisit fixations per trial. To explain why there can be more chunk fixations than objects within both screen sides the definition of a chunk fixations is crucial. A fixation is classified as chunk if it is the first fixation to a not already fixated object within a switch. In the next switch to this side, the same object may be classified as a chunk fixation again. So, despite of the counter intuitive appearance of this finding, we can ascertain that some participants fixated the same objects in more than one switch. Furthermore, we can discern that participants revisited previously fixated objects on average 1.7 times. Participants can be distinguished into two groups: the first group with 10 participants made a nearly equal number of chunk and revisit fixations, meaning for every chunk fixation approximately one revisit fixation was made. Furthermore, the mean number of chunk fixations was not greater than 40 per trial. The other group consisting of only 5 participants made more than 60 revisit fixations and generally more chunk fixations. These two groups can be differentiated best by the bimodality of the mean number of revisit fixations per trial.

One of the most interesting findings regarding the mean number of fixations per trial is the difference between chunk and revisit fixations. While the number of chunk fixations decreased slightly from Block 1 to Block 2, the number of revisits increased by 67%. Similar effects could be seen regarding the complexity: while the mean number of chunk fixations per trial increased slightly by around 5 fixations in complex compared to easy trials, the mean number of revisit fixations increased by nearly 20 fixations per trial in complex trials. While the number of chunk fixations were rather constant, the previously found difference resulted mainly from the increase in revisit fixations. Especially in Block 2 and in complex trials more revisit fixations were used, indicating that revisits are used in order to ensure information can be maintained even though cognitive demands due to perceptual expenses or task specifications are increased.

Since there was a difference between the sides, a more in detail examination accentuated that the number of chunk fixations stayed approximately the same on both sides whereas the number of revisit fixations was considerably larger on the left compared to the right side. Once more, these findings suggest that there is a different mechanism between chunk and revisit fixations in play. As expected, there was nearly no difference between the number of chunk fixations between the sides because each fixated object on the left was compared to the corresponding object on the right. However, in order to keep the objects on the left in memory, those objects had to be refixated.

4.1.3 Total fixation time

Interestingly we saw that the mean total fixation time per trial stayed approximately the same between the Blocks even though the number of fixations per trial and the fixation duration differed between Block 1 and Block 2. The mean total fixation time per switch however differed between the Blocks. In Block 2 the mean total fixation time per switch increased by 75.5%.

The total fixation time per trial and per switch was affected by both the delay and the complexity. In trials with a long delay the total fixation time was longer, suggesting a bigger cognitive demand. For complex objects the same effect was found but on a larger scale. While the difference between a short and long delay was approximately six seconds per trial, the difference between easy and complex objects was nearly 15 seconds, which was nearly half of the total fixation time in easy trials. Both those findings are in correspondence with previous findings from this project as well as the literature (Hardiess et al., 2011, Hardiess et al., 2008a). As expected, we could observe a behavioral adaption to a change in tasks as well as object properties associated with corresponding cost alternations. Once again, we could affirm that a longer delay in which information has to be stored is more demanding compared to a short delay. Furthermore, we could confirm that the use of three object features compared to just two is also more cognitive demanding.

It is to note, that the total fixation time per switch was less effected by the delay or object complexity compared to a trial level. One explanation for this finding is that participants used more switches in complex compared to easy conditions, which would explain the larger total fixation time on a trial level but not on a switch level. Another explanation could be the exclusion of misfixations to the wrong side the beginning of a trial on a switch level but not on a trial level. However, from a total of 600 trials only 87 started with a fixation to the wrong side.

By looking at the interaction between Block and delay regarding the mean total fixation time per trial we can see once again that a short delay by itself without other limitations in Block 1 was doable for participants. However, by including a limitation on the number of switches a participant could make in Block 2, participants fixated longer even in the short delay condition, indicating a greater difficulty. Similarly, we saw that participants mean total fixation time per trial was the longest in trials with a combination of a long delay and complex objects and that the increase in total fixation time was steeper compared to the complex and short combination.

While the interaction between Block and complexity regarding the mean total fixation time was not significant in a trial level, it was significant on the level of a switch. Once again it could be seen that complex objects were more demanding, especially in Block 2.

Participants fixated longer on the left compared to the right side with a decrease in mean total fixation time per trial to 83.3% on the right side. This finding probably stems from the larger number of fixations on the left side, as the fixation duration did not differ between the sides. A comparable effect could be seen on switch level. This corresponds to the findings from Pomplun et al. (2001) were participants exhibited longer total fixation times in the search in contrast to the comparison phase.

Similar to Over et al. (2007) it was found that the mean total fixation time was longer for the first few objects, not just the first as suggested by them. This effect mainly resulted from the larger number of fixations to the first few objects as discussed before. The number of objects in question seems to be similar to the average chunk size. In contrary to Over et al. (2007) we saw no increase in total fixation time for later fixations, assuming that the object number correlates with the fixation number. It is to note that we did not analyze the ordinality of fixations by themselves but use the object number synonymously. Therefore, this deduction might be due to a lack of comparable analyses. It would be interesting to see whether the total fixation time increases with the ordinal fixation number, as Over et al. (2007) found.

By including the differentiation of a fixation in chunk or revisit some new discoveries were made. First of all, as seen in the number of fixations per trial, participants fall into one of two groups. In the first group the mean revisit total fixation time decreases in comparison to the mean chunk total fixation time per trial. In the second group consisting of only 5 participants the opposite was the case.

Apart from that, a cross interaction between the fixation type and the Blocks was found, which explains why the mean fixation duration per trial did not differ between Blocks. While the mean revisit total fixation time per trial increased by 51.1% from Block 1 to Block 2, the mean chunk total fixation time per trial in Block 2 decreased to 77.3% of the duration from Block 1. This finding alone emphasizes the importance of eye tracking data and classification of fixations in chunk and revisit. Another example for the importance is the difference in mean chunk or revisit total fixation time per trial depending on the delay. While the mean revisit total fixation time increased by approximately 2 seconds from a short to a long delay, the mean chunk total fixation time increased double this time from 16.8 to 20.8 seconds. Besides those two new findings, we also saw that the mean revisit total fixation time per trial increased extraordinarily in Block 2 from the easy to the complex condition.

By including the screen side, three more discoveries were made. Firstly, the mean chunk total fixation time per trial stayed approximately the same on both screen sides while the revisit total fixation time on the right was only 74.5 % of the revisit total fixation time on left. Secondly, the mean total fixation time in trials with a short delay decreased the most from the left to the right side. Lastly, the mean revisit total fixation time decreased disproportionately from the left to the right side in complex conditions. Those three findings regarding the total fixation time as well as our previous findings regarding the number of fixations support the hypothesis of different mechanisms on the left and right screen side and of a difference between chunk and revisit fixations.

4.1.4 Number of switches

As expected by our task there was a decrease in hemifield switches between Block 1 and Block 2. Exceeding our expectations this decrease in switches was more than the 75% specified by the task. Participants managed to decrease the number of hemifield switches to nearly 60% of the ones used in Block 1. This even greater reduction highlights the capabilities of participants adapt to new circumstances (Kibbe and Kowler, 2011, Hardiess et al., 2011, Hardiess and Mallot, 2015). This furthermore stresses the assumption that humans usually do not behave on their maximal capacity (Simon, 1955).

In addition to the increase in the number of fixations as well as the total fixation time we saw that participants needed more switches in complex object conditions, intensifying our stance that the object complexity had a large effect on the strategy trade-off. Another explanation for the increase in the number of switches for complex objects could be that participants continued to search until they found another difference or felt sure enough that there is only one. Such a persistence in search behavior was also found by Scinto et al. (1986).

Because participants started on the left side, which was counted as a switch to the left, there were more switches to the left compared to the right hemifield.

As mentioned before there were 87 from 600 trials that started on the right side, even though there was nothing visible at this point. Though it is unlikely, this exclusion of fixations on the level of a switch could be a possible explanation for the difference between the findings on switch on trial level.

4.1.5 Chunk size and number of revisits

First of all, we saw that the chunk size and the number of revisits were different from another. For the in-detail discussion we will first look at the chunk size, then at the number of revisits and lastly at the correlation between both of them.

4.1.5.1 Chunk size

As stated before, chunk size is defined as the number of distinct objects within a switch and therefore resembles working memory capacity. The mean working memory capacity of the 15 participants was approximately 5, which corresponds with the three to five chunks from Cowan (2001) and the magical number seven plus minus two by Miller (1956b). Similar to all other parameters, there was a huge variability between the participants. The average chunk size of participants ranged from around 2.5 to nearly 9 objects, which resembles the variability found in literature (Cowan, 2010). As discussed in the beginning working memory capacity is hard to determine especially when participants use strategies to maximize their performance. Our results lay more on the bigger end of the working memory capacity spectrum.

The chunk size increased by nearly 1.7 objects from Block 1 to Block 2, indicating a behavioral adaption to the new task. On a basal level this means that participants fixated more distinct objects in Block 2 to suffice for the limitation in the number of switches there were allowed to do. This furthermore suggests that participants did not use their maximal capacity in Block 1, leaving room for such an adaption.

The chunk size stayed approximately the same across the different conditions. This was a rather unexpected finding as we assumed that participants would fixate fewer objects if they are more complex or if the delay was long in order to suffice for the increased cognitive demand associated. This assumption was made for a few reasons: first of all, an adaption in chunk size to the task limitations in Block 2 was found, suggesting that behavior can be adjusted according to demands. This finding enables a new perceptive on the behavioral trade-off and its limitations, as we saw that participants did fixate more often, but not new objects. Second, there are studies that showed an increase in used working memory capacity with an increase in memory load. While the difference in chunk size and therefore in working-memory capacity varied between the participants, cognitive demands like delay and feature complexity by themselves did not elicit a change in working-memory capacity use. As we know working memory capacity is limited, however it is likely that humans do not operate on their maximum if not necessary. Our findings, even though unexpected at first, aligns with findings from Chen and Cowan (2009). They found that when verbal reversal was prohibited participants were able to remember three chunks independent of whether the chunks consisted of singletons or learned word pairs. This is very similar to the present study where the objects could either have two or three features, which were learned in the beginning.

4.1.5.2 Number of revisits

By looking at the number of revisits, we saw that participants did on average 7.9 refixations per hemifield switch and 51.8 revisits per trial. According to Gilchrist and Harvey (2000) the number of refixations is an indicator of limited functional memory. In their study they considered the two extremes of memory performance perfect and no memory indicated by the absence or occurrence of revisits. As expected, memory was not perfect because participants refixated previously fixated objects. Nevertheless, they found that there was an inhibition of return to recently fixated objects. This inhibition was only two to three fixations long, after which refixations started to reoccur. In the current study many revisits were made by the participants, suggesting that their memory was not perfect.

Furthermore, we found that the number of revisits in Block 2 more than doubled, which further increases the body of evidence for a behavioral adaptation, elicited by the task. This furthermore aligns with an increase in rehearsal in order to reactivate the items after 10 s to 20 s (Cowan, 1992). Reactivation and consolidation of the stored information enabled the participants to maintain information even though it exceeded the normal working memory span and temporal limitations.

While the delay and complexity by themselves did not affect the mean number of revisits, there was an interaction between Block and complexity. In Block 2 the mean number of revisits increased more for complex objects compared to easy ones. This also indicates that participants found the restrictions in Block 2 along with complex objects the most demanding. It furthermore aligns with the findings from Scinto et al. (1986) and the increase in the number of switches for complex objects, both suggesting that participants kept looking as long as they did not find the differences or were not sure whether there were any more.

Furthermore, the number of revisits was greater on the left side, indicating an increased demand in cognitive resources needed for the acquisition and maintenance of the objects in comparison to retrieval on the right side. The additional refixations on the left side seem to consolidate acquired information, which was not needed on the right side. As suggested by Gould and Dill (1969) refixations are used to obtain more information about the object. I would include that refixations are also used for the maintenance of information not just for more detailed acquisition. In Gould and Dill's (1969) experiment, participants tended to refixate target pattern about twice as many times as compared to non-target patterns. This is something that was not examined here but could be an interesting aspect to look into in further experiments. The aspect of information maintenance is one of the basic mechanisms of working memory suggested Cowan (1999). While encoding and representation would have been the same on both sides, maintenance of information on the left hemifield seems to play an important role in the difference in the number of revisits. On the right hemifield retrieval of previously stored information should come into play.

4.1.5.3 Correlation between chunk size and number of revisits

The Chunk size had to be at least one in order to solve the given task but was limited due to capacity restraints of the working memory. The number of revisits a participant could make was not restricted by itself but due to the task formulation of solving the task as fast and as good as possible an intrinsic limitation could apply. Participants could be motivated to reduce the number of revisits in order to suffice the task demand: solve the task as fast and as good as possible.

With an increase in chunk size the number of revisits increased exponentially, which could be seen either on the level of participants and on the level of a switch independent of participant. These findings also enhance the body of evidence that participants use revisits in order to maintain the encoded information. It has to be noted that the number of revisits was not distributed normally for most chunk sizes. However, by comparing the mean, a corrected mean in which outliers were excluded as well as the median we decided to use the mean for further investigations anyways. The main reason for this decision was that the difference between all three measurements was rather small and we wanted to depict the full behavioral range if possible. Outliers most likely had a reason for their occurrence and probably resembled a present behavior, so they should not be excluded by default.

Even though delay and complexity did not affect either the mean chunk size or the mean number of revisits, we saw an effect regarding the combinations of those two parameters. Especially for complex objects there were more revisits at higher chunk sizes. Similarly, a long delay led to more revisits in higher chunk sizes compared to a short delay.

Because simply looking at the mean does not enable us to make any claims regarding the prevalence of certain chunk sizes and therefore depict the strategy in an adequate manner, we also looked at the number of observations per chunk size. Once again, we could see that in order to suffice for complex conditions, participants fixated more. The second main point that could be seen was that participants fixated two distinct objects per switch most frequently. There were less than 100 observations for chunk sizes larger than five. While the number of observations at a chunk size of 15 was rather small, there was still a difference between easy and complex conditions. The number of observations at a chunk size of 15 in easy conditions was 7, independent of delay. In complex conditions this number was more than tripled with 26 observations.

While this difference may seem small, we can see by looking at the summed number of fixations per chunk size for the four combinations, that this number resembles a total of 1131 revisit fixations for the complex long condition and 940 revisits for the complex short condition. Even though most switches had a chunk size of 2, most revisits in complex trials were done in switches with a chunk size of 4 and 15.

4.1.6 Correlations

The fixation duration decreased with an increase in the number of fixations per trial and per switch between the participants. Different from Vlaskamp et al. (2005) items were always placed in the same spacing each trial, so the correlation of the number of fixations and the mean fixation duration was not cause by a difference in the visual span. Interestingly, in literature fixation duration is only compared to the ordinality of a fixation and not the number of fixations (Over et al., 2007, Antes, 1974). Even tough it would be interesting if this was the first time, to my knowledge, that this was done, it is more likely that this correlation was not fruitful before.

While this might not be as visible with only 15 participants it becomes very clear when each trial for each participant is plotted (Figure 67). There is an accumulation of trials in which participants used fewer and shorter fixations. However, there are some outliers. A strategy that decreases the time spend per trial would fall into the first category: short durations of a fixation and less fixations. It nevertheless is impressive that this could also be shown to an extent on participant level.

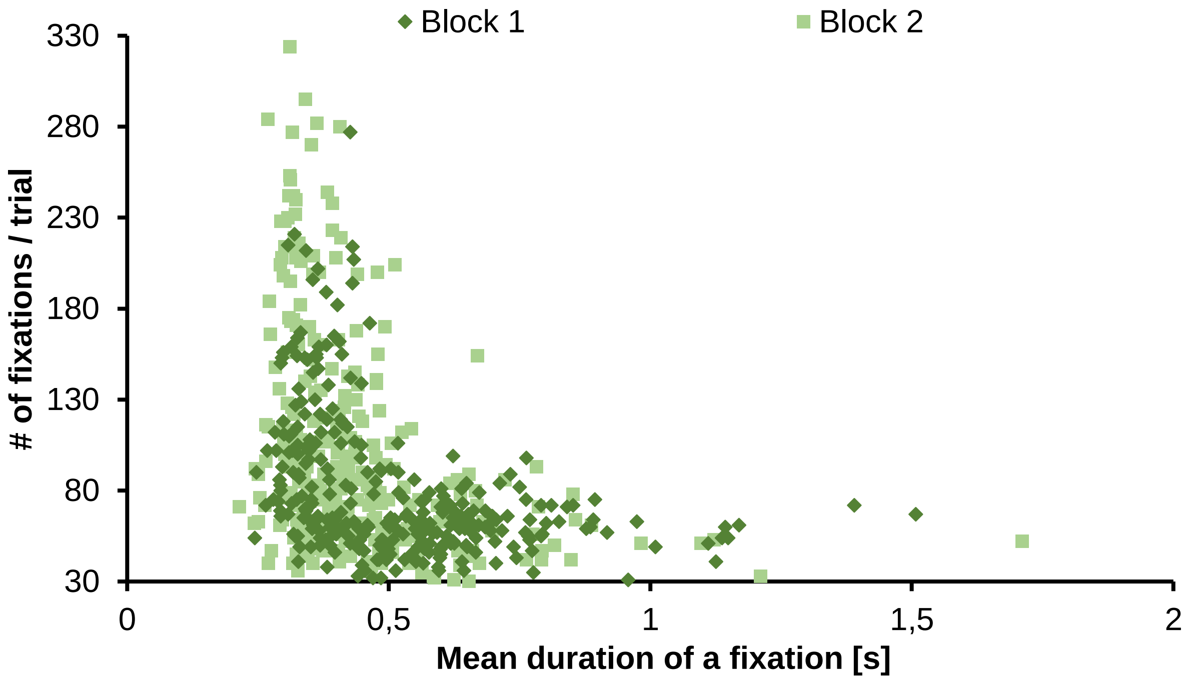

Figure 67: Number of fixations per trial per fixation duration
Number of fixations per trial plotted against the mean fixation duration [s] for both Blocks: Block 1 (dark green) and Block 2 (light green) averaged over all fixations of a trial. Every data point is a single trial of a participant.

Accordingly, chunk size was correlated with the duration of a chunk fixation, while the mean duration of a revisit fixation depended on the number of revisits. Egan and Schwartz (1979) also found that the size of a chunk increased with additional study time. This finding is similar to findings from Meghanathan et al. (2015) in which the fixation duration showed to be a good predictor for memory load.

The number of hemifield switches was correlated with first the number of fixations per switch and second the total fixation time per switch. According to work from Hardiess and Mallot (2015) the memorization acquisition trade-off in the comparative visual search paradigm can be described by processing time and number of gaze shifts. However, they approximated processing time by the time between two mouse clicks to between two switches. Fixation duration is a more accurate and direct measure of processing time and may explain the difference in power between their and the present results. As expected, due to the correlation with the number of fixations per switch, an increase in chunk size and number of revisits also decrease the number of switches participants used. Interestingly, the chunk size has shown to be the better predictor for switches in comparison to the number of revisits or total number of fixations. Therefor is could be suggested that chunk size would be a better estimator for working memory strategy use.

4.2 General discussion

After an in-depth result discussion, the hypotheses are examined and compared with literature.

It was assumed that there will be a strategy change between the first and the second Block. This strategy switch was found mainly in the number of switches, total fixation time and number of switches. The findings are similar to behavioral findings acquired in my bachelor thesis (Bräutigam, 2018), and in findings from Hardiess et al. (2011) and Hardiess et al. (2008b). While the trade-off between acquisition resembled by the number of switches and memorization resembled by the fixation time per switch could also be found by using eye tracking data, it was worse compared to behavioral approaches. This difference could result from the fact that fixation duration and therefor fixation time per switch is a direct measure of cognitive processing. Time per switch on the other hand is an indirect measure used to approximate this processing time, but it also includes non-task or processing related durations. While the acquisition memorization strategy trade-off by looking at processing time measured directly by fixation duration got weaker new strategies were discovered. It seems that chunk size as a predictor of cognitive load or information contend and therefore memorization strategy is better suited to use in this strategy trade-off. Participants furthermore not only modulated their fixation duration and the number of fixations in order to fulfill the new task demands to solve each trial as fast and as good as possible. It appears that there is another trade-off participant have to make to increase the number of objects they want to compare. By increasing the number of chunk fixations the number of revisits increased to. Similar to Gould and Dill (1969) and Eriksson et al. (2015) refixations are used to obtain more information about the object and counteract working memory constraints. It is unclear if this maintenance of information is the result of a reactivation, a spatial rehearsal process or simply due to the reallocation of attention to these items.

As set by the task requirements, there were less hemifield switches in Block 2. The extent of this reduction however was rather unexpected as it exceeded the requirements by the task substantial. This finding suggests that participants did not perform on their maximum and had room for improvement if needed. A similar effect was found by Simon (1955). It seems that the presence of the switch countdown led to participants using less switches. This might be the result of an additional intrinsic motivation participants acquired by seeing this limitation. As hypothesized before participants counteracted the limitation of hemifield switches by changing their acquisition behavior, e.g. making more but shorter fixations in Block 2, which interestingly balanced out in the mean total fixation time.

Furthermore, it was predicted that trials with complex objects and a long delay will be harder for participants to solve. This was visible in the longer mean fixation duration and longer total fixation time for both the long delay and the complex objects as well as in the higher number of fixations and number of switches in trials with complex objects. This challenge seemed to increase even more in Block 2 as seen in the interactions regarding the mean fixation duration between Block and delay or complexity. In the same way the interaction between Block and complexity regarding the number of fixations as well as the number of switches and lastly the interaction between Block and delay in the total fixation time increased in Block 2. Furthermore, the combination of a long delay duration and complex objects seemed more taxing as well, as seen in the number of fixations and total fixation time. In general, it can be deduced that while delay had a slight impact on the strategy, object complexity had the bigger effect, which can most likely be attributed to the associated increase in perceptual and cognitive demands. An example for such a behavioral modulation in a comparative visual search with increasing difficulty was seen in Hardiess and Mallot (2015). They also found that delay had a smaller effect on error rates and response time compared to object complexity. In their experiment they used color as a simple feature and silhouettes of animals as complex objects and proved that trials with complex objects were harder for participants to solve. In another experiment gap size is varied in tasks where the letter O has to be found amongst Cs (Hooge and Erkelens, 1996, Vlaskamp et al., 2005). It has been shown that reducing the gap size increases the perceptual difficulty of a target. By introducing three features every object had both colors grey and black which makes the targets more similar. As investigated previously by Vlaskamp et al. (2005) both the similarity and the spacing would affect search time and be indicative of difficulty. Similar to Hardiess and Mallot (2015), Hardiess et al. (2011) and Droll and Hayhoe (2007) is rather stable within a subject but differs vastly between subjects.

Additionally, it a difference between the two hemifields was found that presumably can be attributed to the difference in cognitive processes involved. This difference was most like due to the nature of this task as it started on the left side. Participants used the left side as a template, the right side was compared to. The main difference in cognitive processes is therefore that the information obtained on the left side is maintained on the right side for comparing the information on the right side. This hypothesis could be confirmed because especially the number of fixations and as a result also the total fixation duration was larger on the left side. By including the differentiation between chunk and revisit fixation it became obvious that this difference stemmed from an increase in refixations on the left side compared to the right. The basic mechanisms of working memory proposed by Cowan (1999) can help to understand why revisits were done. While items had to be encoded and represented on both hemifields, the maintenance of information was only necessary on the left side and for the delay duration. According to Cowan (1999) information is lost after around 10 to 20 seconds if there was no reactivation. In order to maintain this information participants revisited the previously fixated objects to refocus their attention on these items. Because verbal rehearsal was prohibited by using nonsense words participants had to use different methods to maintain the information (Cowan, 2010). It cannot be excluded that participants used non-verbal rehearsal or externalized information by using their finger to count, however it seems that the main strategies used to maintain a representation was revisiting the objects and therefore fortifying the representation.

Lastly, we anticipated that the classification of fixations as chunk or revisit will lead to new insights regarding cognitive processes within a comparative visual search task. Chunk size and number of revisits alone enabled some new insights on the modulation of search behavior. Especially in regard to the understanding of revisits as a strategy used to maintain information, this classification proved very substantial. It is most likely that this difference occurred due to maintenance of information on the left compared to the right side (Cowan, 1999, Braisby and Gellatly, 2012, Eriksson et al., 2015). The use of refixations is suggested to counteract working memory capacity limitations (Eriksson et al., 2015). Furthermore, the chunk fixations were longer than revisit fixations, indicating a higher processing demand. The first fixation of an object within a switch is when the representation of this object is formed. A refixation of this object simply serves as a renewal of the already present information. Especially in more demanding trials like in Block 2, with complex features or a long delay the differentiation between chunk and revisit fixation highlighted new findings. In trials with a longer delay participants chunk fixation duration increased, while revisit fixation duration stayed the same. In Block 2 the number of revisit fixations increased drastically, while the number of chunk fixations stayed the same. These effects became even more obvious when looking at complex and easy trials. Especially the investigation of total fixation time by inclusion of fixation type showed why there was no difference between Block 1 and Block 2 in the previous superficial analysis. All results considered it is needless to say that the differentiation between chunk and revisit fixations led to some very interesting new findings and enable further research especially regarding the difference in processing.

In conclusion all hypothesis could be confirmed and bring forth new insights as well as further questions to be studied in the future.

4.3 Future Research

Based on the present findings the next step is to conduct more experiments using the differentiation between chunk and revisit fixations. These experiments should opt to answer the question whether chunk size might really be a better operationalization of memorization strategy compared to processing time. Furthermore, it would be important to investigate if the difference between revisit and chunk fixations really stems from differences in processing as discussed before or if they are the result of some other underlying factor that was not detected yet.

Going deeper into hypotheses regarding working memory limitations it would be interesting to design an experiment tackling the question if working memory is a flexible resource or is better explained by a slot model. This could also be done by including the differentiation between chunk and revisit fixations.

Regarding the forced decrease in the number of switches and the observed behavioral modulation exceeding the requirements it would be interesting to see how much further such limitations can be extended, before humans are unable to perform the task. In this context a gradual increase in delay and feature complexity by varying color, shape and other factors seem advantageous for advancement in the field of strategy adaption and behavioral modulation. Furthermore, it seems possible that the knowledge about certain limitations alone is enough to make participants “try harder”. So, changing the task itself might be viable to investigate this aspect of human cognitive capabilities. The comparative visual search approach seems very promising to address those questions.

Research regarding cognitive processing, behavioral adaptation and strategy modulation will lead to important discoveries concerning human cognition and will inevitably deepen our understanding of the Blackbox that is our brain. Even though research regarding human cognition and understanding of behavior has come a long way, there are still many more questions to be answered. Some of those question will be addressed and hopefully answered in my dissertation.

VII. Literature

ALLEN, G. L. & HAUN, D. B. 2004. Proximity and precision in spatial memory. *Human spatial memory: Remembering where*, 41-63.

ALLON, A. S. & LURIA, R. 2017. Compensation mechanisms that improve distractor filtering are short-lived. *Cognition,* 164**,** 74-86.

ALLON, A. S., VIXMAN, G. & LURIA, R. 2019. Gestalt grouping cues can improve filtering performance in visual working memory. *Psychological research,* 83**,** 1656-1672.

ANTES, J. R. 1974. The time course of picture viewing. *Journal of experimental psychology,* 103**,** 62.

ARNELL, K. M. & STUBITZ, S. M. 2010. Attentional blink magnitude is predicted by the ability to keep irrelevant material out of working memory. *Psychological Research,* 74**,** 457-467.

ASHBY, N. J., JOHNSON, J. G., KRAJBICH, I. & WEDEL, M. 2016. Applications and innovations of eye-movement research in judgment and decision making.

ATKINSON, R. C. & SHIFFRIN, R. M. 1968. Human memory: A proposed system and its control processes. *Psychology of learning and motivation,* 2**,** 89-195.

AUDIBERT, J.-Y., MUNOS, R. & SZEPESVÁRI, C. 2009. Exploration–exploitation tradeoff using variance estimates in multi-armed bandits. *Theoretical Computer Science,* 410**,** 1876-1902.

AWH, E., JONIDES, J. & REUTER-LORENZ, P. A. 1998. Rehearsal in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance,* 24**,** 780.

AWH, E., JONIDES, J., SMITH, E. E., BUXTON, R. B., FRANK, L. R., LOVE, T., WONG, E. C. & GMEINDL, L. 1999. Rehearsal in spatial working memory: Evidence from neuroimaging. *Psychological Science,* 10**,** 433-437.

AWH, E., JONIDES, J., SMITH, E. E., SCHUMACHER, E. H., KOEPPE, R. A. & KATZ, S. 1996. Dissociation of storage and rehearsal in verbal working memory: Evidence from positron emission tomography. *Psychological Science,* 7**,** 25-31.

BACON, W. F. & EGETH, H. E. 1997. Goal-directed guidance of attention: evidence from conjunctive visual search. *Journal of Experimental Psychology: Human Perception and Performance,* 23**,** 948.

BADDELEY, A. 1986. Working memory Oxford: Oxford University Press. Clarendon Press.

BADDELEY, A. 1992. Working memory. *Science,* 255**,** 556-559.

BADDELEY, A. 2000. The episodic buffer: a new component of working memory? *Trends in cognitive sciences,* 4**,** 417-423.

BADDELEY, A. 2003. Working memory and language: An overview. *Journal of communication disorders,* 36**,** 189-208.

BADDELEY, A. 2007. *Working memory, thought, and action*, OuP Oxford.

BADDELEY, A. 2010. Working memory. *Current biology,* 20**,** R136-R140.

BADDELEY, A., LEWIS, V. & VALLAR, G. 1984. Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology Section A,* 36**,** 233-252.

BADDELEY, A. D. & HITCH, G. 1974. Working memory. *Psychology of learning and motivation.* Elsevier.

BADDELEY, A. D., HITCH, G. J. & ALLEN, R. J. 2009. Working memory and binding in sentence recall. *Journal of Memory and Language,* 61**,** 438-456.

BALLARD, D. H., HAYHOE, M. M. & PELZ, J. B. 1995. Memory representations in natural tasks. *Journal of cognitive neuroscience,* 7**,** 66-80.

BARROUILLET, P., BERNARDIN, S. & CAMOS, V. 2004. Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General,* 133**,** 83.

BARROUILLET, P. & CAMOS, V. 2012. As time goes by: Temporal constraints in working memory. *Current Directions in Psychological Science,* 21**,** 413-419.

BARROUILLET, P., DE PAEPE, A. & LANGEROCK, N. 2012. Time causes forgetting from working memory. *Psychonomic Bulletin & Review,* 19**,** 87-92.

BARROUILLET, P., PORTRAT, S., VERGAUWE, E., DIEPENDAELE, K. & CAMOS, V. 2011. Further evidence for temporal decay in working memory: reply to Lewandowsky and Oberauer (2009).

BASDEN, B. H., BASDEN, D. R. & GARGANO, G. J. 1993. Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 19**,** 603.

BAYLIS, G. C. & DRIVER, J. 1992. Visual parsing and response competition: The effect of grouping factors. *Perception & Psychophysics,* 51**,** 145-162.

BOOT, W. R., BECIC, E. & KRAMER, A. F. 2009. Stable individual differences in search strategy?: The effect of task demands and motivational factors on scanning strategy in visual search. *Journal of Vision,* 9**,** 7-7.

BRAINARD, D. H. 1997. The psychophysics toolbox. *Spatial vision,* 10**,** 433-436.

BRAISBY, N. & GELLATLY, A. 2012. *Cognitive psychology*, Oxford University Press.

BRÄUTIGAM, L. 2018. Individuelle, strategieabhängige Beeinflussung des Trade-offs zwischen Akquisition und Memorisierung im Paradigma der visuellen Suche. *Bachelorarbeit der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen*.

BROWN, J. 1958. Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology,* 10**,** 12-21.

BRUCE, N. D. & TSOTSOS, J. K. 2009. Saliency, attention, and visual search: An information theoretic approach. *Journal of vision,* 9**,** 5-5.

BUTKO, N. J. & MOVELLAN, J. R. I-POMDP: An infomax model of eye movement. 2008 7th IEEE International Conference on Development and Learning, 2008. IEEE, 139-144.

CAI, D., HE, X., LI, Z., MA, W.-Y. & WEN, J.-R. Hierarchical clustering of WWW image search results using visual, textual and link information. Proceedings of the 12th annual ACM international conference on Multimedia, 2004. 952-959.

CAPALDI, E., NAWROCKI, T. M., MILLER, D. J. & VERRY, D. R. 1986. Grouping, chunking, memory, and learning. *The Quarterly Journal of Experimental Psychology Section B,* 38**,** 53-80.

CASTELHANO, M. S. & HENDERSON, J. M. 2008. Stable individual differences across images in human saccadic eye movements. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale,* 62**,** 1.

CHASE, W. G. & SIMON, H. A. 1973. Perception in chess. *Cognitive psychology,* 4**,** 55-81.

CHEN, S., KOCSIS, A., LIESEFELD, H. R., MÜLLER, H. J. & CONCI, M. 2020. Object-based grouping benefits without integrated feature representations in visual working memory. *Attention, Perception, & Psychophysics***,** 1-18.

CHEN, Z. & COWAN, N. 2009. Core verbal working-memory capacity: The limit in words retained without covert articulation. *Quarterly journal of experimental psychology,* 62**,** 1420-1429.

COHEN, A. & IVRY, R. B. 1991. Density effects in conjunction search: evidence for a coarse location mechanism of feature integration. *J Exp Psychol Hum Percept Perform,* 17**,** 891-901.

COHEN, J. D., MCCLURE, S. M. & YU, A. J. 2007. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 362**,** 933-942.

COWAN, N. 1992. Verbal memory span and the timing of spoken recall. *Journal of Memory and Language,* 31**,** 668-684.

COWAN, N. 1999. An embedded-processes model of working memory. *Models of working memory: Mechanisms of active maintenance and executive control,* 20**,** 506.

COWAN, N. 2001. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and brain sciences,* 24**,** 87-114.

COWAN, N. 2008. What are the differences between long-term, short-term, and working memory? *Progress in brain research,* 169**,** 323-338.

COWAN, N. 2010. The magical mystery four: How is working memory capacity limited, and why? *Current directions in psychological science,* 19**,** 51-57.

COWAN, N. 2016. *Working memory capacity: Classic edition*, Psychology press.

COWAN, N. & MOREY, C. C. 2006. Visual working memory depends on attentional filtering. *Trends in cognitive sciences,* 10**,** 139-141.

DAEHLER, M. W., HOROWITZ, A. B., WYNNS, F. C. & FLAVELL, J. H. 1969. Verbal and nonverbal rehearsal in children's recall. *Child Development***,** 443-452.

DE GROOT, A. D. 1946. *Het denken van den schaker: een experimenteel-psychologische studie*, Noord-Hollandsche Uitgevers Maatschappij Amsterdam.

DE LILLO, C. 2004. Imposing structure on a Corsi-type task: Evidence for hierarchical organisation based on spatial proximity in serial-spatial memory. *Brain and Cognition,* 55**,** 415-426.

DEL GIUDICE, M. & CRESPI, B. J. 2018. Basic functional trade-offs in cognition: An integrative framework. *Cognition,* 179**,** 56-70.

DI GIACOMO, E., DIDIMO, W., GRILLI, L., LIOTTA, G. & PALLADINO, P. Whatsonweb+: An enhanced visual search clustering engine. 2008 IEEE Pacific Visualization Symposium, 2008. IEEE, 167-174.

DOLAN, R. J. & DAYAN, P. 2013. Goals and habits in the brain. *Neuron,* 80**,** 312-25.

DREW, T., BOETTCHER, S. E. & WOLFE, J. M. 2017. One visual search, many memory searches: An eye-tracking investigation of hybrid search. *Journal of vision,* 17**,** 5-5.

DROLL, J. A. & HAYHOE, M. M. 2007. Trade-offs between gaze and working memory use. *Journal of Experimental Psychology: Human Perception and Performance,* 33**,** 1352.

DUNLOSKY, J., BAILEY, H. & HERTZOG, C. 2011. Memory enhancement strategies: What works best for obtaining memory goals? *Enhancing Cognitive Fitness in Adults.* Springer.

EGAN, D. E. & SCHWARTZ, B. J. 1979. Chunking in recall of symbolic drawings. *Memory & Cognition,* 7**,** 149-158.

ENGBERT, R., LONGTIN, A. & KLIEGL, R. 2002. A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision research,* 42**,** 621-636.

ENGLE, R. W. & KANE, M. J. 2004. Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychology of learning and motivation,* 44**,** 145-200.

ERIKSSON, J., VOGEL, E. K., LANSNER, A., BERGSTRÖM, F. & NYBERG, L. 2015. Neurocognitive architecture of working memory. *Neuron,* 88**,** 33-46.

FINDLAY, J. M., FINDLAY, J. M. & GILCHRIST, I. D. 2003. *Active vision: The psychology of looking and seeing*, Oxford University Press.

FINDLAY, J. M. & WALKER, R. 1999. A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences,* 22**,** 661-674.

FLAVELL, J. H., BEACH, D. R. & CHINSKY, J. M. 1966. Spontaneous verbal rehearsal in a memory task as a function of age. *Child development***,** 283-299.

FRANCO-WATKINS, A. M. & JOHNSON, J. G. 2011. Decision moving window: Using interactive eye tracking to examine decision processes. *Behavior research methods,* 43**,** 853.

FUKUDA, K. & VOGEL, E. K. 2009. Human variation in overriding attentional capture. *Journal of Neuroscience,* 29**,** 8726-8733.

GALPIN, A. J. & UNDERWOOD, G. 2005. Eye movements during search and detection in comparative visual search. *Perception & Psychophysics,* 67**,** 1313-1331.

GIDLÖF, K., WALLIN, A., DEWHURST, R. & HOLMQVIST, K. 2013. Using eye tracking to trace a cognitive process: Gaze behaviour during decision making in a natural environment.

GILCHRIST, I. D. & HARVEY, M. 2000. Refixation frequency and memory mechanisms in visual search. *Current Biology,* 10**,** 1209-1212.

GIROD, B., CHANDRASEKHAR, V., CHEN, D. M., CHEUNG, N.-M., GRZESZCZUK, R., REZNIK, Y., TAKACS, G., TSAI, S. S. & VEDANTHAM, R. 2011. Mobile visual search. *IEEE signal processing magazine,* 28**,** 61-76.

GOBET, F., LANE, P. C., CROKER, S., CHENG, P. C., JONES, G., OLIVER, I. & PINE, J. M. 2001. Chunking mechanisms in human learning. *Trends in cognitive sciences,* 5**,** 236-243.

GOULD, J. D. 1973. Eye movements during visual search and memory search. *Journal of experimental psychology,* 98**,** 184.

GOULD, J. D. & DILL, A. B. 1969. Eye-movement parameters and pattern discrimination. *Perception & Psychophysics,* 6**,** 311-320.

GRONER, M. T., GRONER, R. & VON MÜHLENEN, A. 2008. The effect of spatial frequency content on parameters of eye movements. *Psychological Research,* 72**,** 601-608.

HARDIESS, G., BASTEN, K. & MALLOT, H. A. 2011. Acquisition vs. memorization trade-offs are modulated by walking distance and pattern complexity in a large-scale copying paradigm. *Plos one,* 6**,** e18494.

HARDIESS, G., GILLNER, S. & MALLOT, H. A. 2008a. Head and eye movements and the role of memory limitations in a visual search paradigm. *J Vis,* 8**,** 7 1-13.

HARDIESS, G., GILLNER, S. & MALLOT, H. A. 2008b. Head and eye movements and the role of memory limitations in a visual search paradigm. *Journal of Vision,* 8**,** 7-7.

HARDIESS, G. & MALLOT, H. A. 2015. Allocation of cognitive resources in comparative visual search–Individual and task dependent effects. *Vision Research,* 113**,** 71-77.

HAYHOE, M. M., SHRIVASTAVA, A., MRUCZEK, R. & PELZ, J. B. 2003. Visual memory and motor planning in a natural task. *J Vis,* 3**,** 49-63.

HENDERSON, J. M., CHANCEAUX, M. & SMITH, T. J. 2009. The influence of clutter on real-world scene search: Evidence from search efficiency and eye movements. *Journal of vision,* 9**,** 32-32.

HENDERSON, J. M., WEEKS JR, P. A. & HOLLINGWORTH, A. 1999. The effects of semantic consistency on eye movements during complex scene viewing. *Journal of experimental psychology: Human perception and performance,* 25**,** 210.

HOFFMAN, J. E. 1998. Visual attention and eye movements. *Attention,* 31**,** 119-153.

HOOGE, I. T. C. & ERKELENS, C. J. 1996. Control of fixation duration in a simple search task. *Perception & Psychophysics,* 58**,** 969-976.

HOOGE, I. T. C. & ERKELENS, C. J. 1998. Adjustment of fixation duration in visual search. *Vision research,* 38**,** 1295-IN4.

HOOGE, I. T. C., VLASKAMP, B. N. & OVER, E. A. 2007. Saccadic search: On the duration of a fixation. *Eye movements.* Elsevier.

HOURIHAN, K. L., OZUBKO, J. D. & MACLEOD, C. M. 2009. Directed forgetting of visual symbols: Evidence for nonverbal selective rehearsal. *Memory & Cognition,* 37**,** 1059-1068.

HWANG, G., JACOBS, J., GELLER, A., DANKER, J., SEKULER, R. & KAHANA, M. J. 2005. EEG correlates of verbal and nonverbal working memory. *Behavioral and Brain Functions,* 1**,** 20.

INAMDAR, S. & POMPLUN, M. Comparative search reveals the tradeoff between eye movements and working memory use in visual tasks. Proceedings of the Annual Meeting of the Cognitive Science Society, 2003.

JOST, K., BRYCK, R. L., VOGEL, E. K. & MAYR, U. 2011. Are old adults just like low working memory young adults? Filtering efficiency and age differences in visual working memory. *Cerebral cortex,* 21**,** 1147-1154.

KAHNEMAN, D. & HENIK, A. 1977. Effects of visual grouping on immediate recall and selective attention. *Attention and performance VI***,** 307-332.

KALANTIDIS, Y., KENNEDY, L., NGUYEN, H., MELLINA, C. & SHAMMA, D. A. Loh and behold: Web-scale visual search, recommendation and clustering using locally optimized hashing. European conference on computer vision, 2016. Springer, 702-718.

KANE, M. J., BLECKLEY, M. K., CONWAY, A. R. & ENGLE, R. W. 2001. A controlled-attention view of working-memory capacity. *Journal of experimental psychology: General,* 130**,** 169.

KANE, M. J. & ENGLE, R. W. 2003. Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of experimental psychology: General,* 132**,** 47.

KATONA, G. 1940. Organizing and memorizing: studies in the psychology of learning and teaching.

KEENEY, T. J., CANNIZZO, S. R. & FLAVELL, J. H. 1967. Spontaneous and induced verbal rehearsal in a recall task. *Child Development*, 953-966.

KENSINGER, E. A. & CORKIN, S. 2003. Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & cognition,* 31, 1169-1180.

KIBBE, M. M. & KOWLER, E. 2011. Visual search for category sets: Tradeoffs between exploration and memory. *Journal of vision,* 11, 14-14.

KIMCHI, R., YESHURUN, Y. & COHEN-SAVRANSKY, A. 2007. Automatic, stimulus-driven attentional capture by objecthood. *Psychonomic Bulletin & Review,* 14, 166-172.

KIMCHI, R., YESHURUN, Y., SPEHAR, B. & PIRKNER, Y. 2016. Perceptual organization, visual attention, and objecthood. *Vision Research,* 126, 34-51.

KIRCHNER, D. M. & KLATZKY, R. L. 1985. Verbal rehearsal and memory in language-disordered children. *Journal of Speech, Language, and Hearing Research,* 28, 556-565.

KIT, D., KATZ, L., SULLIVAN, B., SNYDER, K., BALLARD, D. & HAYHOE, M. 2014. Eye movements, visual search and scene memory, in an immersive virtual environment. *PLoS One,* 9, e94362.

KONG, X., SCHUNN, C. D. & WALLSTROM, G. L. 2010. High regularities in eye-movement patterns reveal the dynamics of the visual working memory allocation mechanism. *Cogn Sci,* 34, 322-37.

KUBOVY, M., HOLCOMBE, A. O. & WAGEMANS, J. 1998. On the lawfulness of grouping by proximity. *Cognitive psychology,* 35, 71-98.

LEWANDOWSKY, S. & OBERAUER, K. 2009. No evidence for temporal decay in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 35, 1545.

LEWANDOWSKY, S., OBERAUER, K. & BROWN, G. D. 2009. No temporal decay in verbal short-term memory. *Trends in cognitive sciences,* 13, 120-126.

LI, C.-H., HE, X., WANG, Y.-J., HU, Z. & GUO, C.-Y. 2017. Visual working memory capacity can be increased by training on distractor filtering efficiency. *Frontiers in psychology,* 8, 196.

LI, C. L., AIVAR, M. P., TONG, M. H. & HAYHOE, M. M. 2018a. Memory shapes visual search strategies in large-scale environments. *Sci Rep,* 8, 4324.

LI, J., QIAN, J. & LIANG, F. 2018b. Evidence for the beneficial effect of perceptual grouping on visual working memory: an empirical study on illusory contour and a meta-analytic study. *Scientific reports,* 8, 1-20.

LIN, Z. & BRANDT, J. W. 2014. Methods and apparatus for visual search. Google Patents.

LOFTUS, G. R. 1985. Picture perception: Effects of luminance on available information and information-extraction rate. *Journal of Experimental Psychology: General,* 114, 342.

LONGONI, A. M., RICHARDSON, J. T. & AIELLO, A. 1993. Articulatory rehearsal and phonological storage in working memory. *Memory & Cognition,* 21, 11-22.

MACKAY, D. G. 1982. The problems of flexibility, fluency, and speed–accuracy trade-off in skilled behavior. *Psychological Review,* 89, 483.

MACREADY, W. G. & WOLPERT, D. H. 1998. Bandit problems and the exploration/exploitation tradeoff. *IEEE Transactions on evolutionary computation,* 2**,** 2-22.

MANDLER, G. 1967. Organization and memory. *Psychology of learning and motivation.* Elsevier.

MANOR, B. R. & GORDON, E. 2003. Defining the temporal threshold for ocular fixation in free-viewing visuocognitive tasks. *Journal of neuroscience methods,* 128**,** 85-93.

MCCONKIE, G., ZOLA, D., GRIMES, J., KERR, P., BRYANT, N., WOLFF, P. & STEIN, J. 1991. Vision and visual dyslexia.

MCNAB, F. & KLINGBERG, T. 2008. Prefrontal cortex and basal ganglia control access to working memory. *Nature neuroscience,* 11**,** 103-107.

MEGHANATHAN, R. N., VAN LEEUWEN, C. & NIKOLAEV, A. R. 2015. Fixation duration surpasses pupil size as a measure of memory load in free viewing. *Frontiers in human neuroscience,* 8**,** 1063.

MEHLHORN, K., NEWELL, B. R., TODD, P. M., LEE, M. D., MORGAN, K., BRAITHWAITE, V. A., HAUSMANN, D., FIEDLER, K. & GONZALEZ, C. 2015. Unpacking the exploration–exploitation tradeoff: A synthesis of human and animal literatures. *Decision,* 2**,** 191.

MENON, G. 1993. The effects of accessibility of information in memory on judgments of behavioral frequencies. *Journal of Consumer Research,* 20**,** 431-440.

MILLER, G. A. 1956a. The magic number seven plus or minus two: Some limits on our capacity for processing information. *Psychological review,* 63**,** 91-97.

MILLER, G. A. 1956b. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review,* 63**,** 81.

MILLER, G. A., GALANTER, E. & PRIBRAM, K. H. 1960. Plans and the structure of behavior.

MIYAKE, A., FRIEDMAN, N. P., EMERSON, M. J., WITZKI, A. H., HOWERTER, A. & WAGER, T. D. 2000. The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive psychology,* 41**,** 49-100.

MOBBS, D., TRIMMER, P. C., BLUMSTEIN, D. T. & DAYAN, P. 2018. Foraging for foundations in decision neuroscience: insights from ethology. *Nat Rev Neurosci,* 19**,** 419-427.

MOFFITT, K. 1980. Evaluation of the fixation duration in visual search. *Perception & Psychophysics,* 27**,** 370-372.

MOFFITT, K. W. 1978. *Eye movements as a correlate of information processing in a visual-search task.* New Mexico State University.

NAJEMNIK, J. & GEISLER, W. S. 2005. Optimal eye movement strategies in visual search. *Nature,* 434**,** 387-391.

OBERAUER, K., FARRELL, S., JARROLD, C. & LEWANDOWSKY, S. 2016. What limits working memory capacity? *Psychological bulletin,* 142**,** 758.

OVER, E., HOOGE, I., VLASKAMP, B. & ERKELENS, C. 2007. Coarse-to-fine eye movement strategy in visual search. *Vision Research,* 47**,** 2272-2280.

PEARSON, D. G., BALL, K. & SMITH, D. T. 2014. Oculomotor preparation as a rehearsal mechanism in spatial working memory. *Cognition,* 132**,** 416-428.

PETERSON, L. & PETERSON, M. J. 1959. Short-term retention of individual verbal items. *Journal of experimental psychology,* 58**,** 193.

POMPLUN, M., SICHELSCHMIDT, L., WAGNER, K., CLERMONT, T., RICKHEIT, G. & RITTER, H. 2001. Comparative visual search: A difference that makes a difference. *Cognitive Science,* 25**,** 3-36.

PORTRAT, S., BARROUILLET, P. & CAMOS, V. 2008. Time-related decay or interference-based forgetting in working memory? *Journal of experimental psychology: Learning, memory, and cognition,* 34**,** 1561.

QUINLAN, P. T. & COHEN, D. J. 2012. Grouping and binding in visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 38**,** 1432.

RAYNER, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin,* 124**,** 372.

RAYNER, K., LI, X., WILLIAMS, C. C., CAVE, K. R. & WELL, A. D. 2007. Eye movements during information processing tasks: Individual differences and cultural effects. *Vision research,* 47**,** 2714-2726.

RAYNER, K., YANG, J., CASTELHANO, M. S. & LIVERSEDGE, S. P. 2011. Eye movements of older and younger readers when reading disappearing text. *Psychology and aging,* 26**,** 214.

ROBISON, M. K., MILLER, A. L. & UNSWORTH, N. 2018. Individual differences in working memory capacity and filtering. *Journal of Experimental Psychology: Human Perception and Performance,* 44**,** 1038.

ROCK, I. & BROSGOLE, L. 1964. Grouping based on phenomenal proximity. *Journal of Experimental Psychology,* 67**,** 531.

SALTHOUSE, T. A. 1979. Adult age and the speed-accuracy trade-off. *Ergonomics,* 22**,** 811-821.

SALTHOUSE, T. A. & ELLIS, C. L. 1980. Determinants of eye-fixation duration. *The American journal of psychology***,** 207-234.

SALTHOUSE, T. A., ELLIS, C. L., DIENER, D. C. & SOMBERG, B. L. 1981. Stimulus processing during eye fixations. *Journal of Experimental Psychology: Human Perception and Performance,* 7**,** 611.

SCHMIDT, J. & ZELINSKY, G. J. 2011. Visual search guidance is best after a short delay. *Vision Research,* 51**,** 535-545.

SCHNEIDER, W. & SHIFFRIN, R. M. 1977. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological review,* 84**,** 1.

SCINTO, L. F., PILLALAMARRI, R. & KARSH, R. 1986. Cognitive strategies for visual search. *Acta psychologica,* 62**,** 263-292.

SHI, S. W., WEDEL, M. & PIETERS, F. 2013. Information acquisition during online decision making: A model-based exploration using eye-tracking data. *Management Science,* 59**,** 1009-1026.

SIMON, H. A. 1955. A behavioral model of rational choice. *The quarterly journal of economics,* 69**,** 99-118.

SIMON, H. A. & GILMARTIN, K. 1973. A simulation of memory for chess positions. *Cognitive psychology,* 5**,** 29-46.

SIVIC, J. & ZISSERMAN, A. 2006. Video Google: Efficient visual search of videos. *Toward category-level object recognition.* Springer.

SMYTH, M. M. 1996. Interference with rehearsal in spatial working memory in the absence of eye movements. *The Quarterly Journal of Experimental Psychology: Section A,* 49**,** 940-949.

STAUB, A. & BENATAR, A. 2013. Individual differences in fixation duration distributions in reading. *Psychonomic Bulletin & Review,* 20**,** 1304-1311.

THEEUWES, J. 1996. Parallel search for a conjunction of color and orientation: The effect of spatial proximity. *Acta Psychologica,* 94**,** 291-307.

TORGESEN, J. & GOLDMAN, T. 1977. Verbal rehearsal and short-term memory in reading-disabled children. *Child Development***,** 56-60.

TREISMAN, A. 1982. Perceptual grouping and attention in visual search for features and for objects. *Journal of experimental psychology: human perception and performance,* 8**,** 194.

TREISMAN, A. M. & GELADE, G. 1980. A feature-integration theory of attention. *Cognitive psychology,* 12**,** 97-136.

TREMBLAY, S., SAINT-AUBIN, J. & JALBERT, A. 2006. Rehearsal in serial memory for visual-spatial information: Evidence from eye movements. *Psychonomic bulletin & review,* 13**,** 452-457.

TRUKENBROD, H. A. & ENGBERT, R. 2014. ICAT: A computational model for the adaptive control of fixation durations. *Psychonomic bulletin & review,* 21**,** 907-934.

VAN DER STIGCHEL, S., BETHLEHEM, R. A., KLEIN, B. P., BERENDSCHOT, T. T., NIJBOER, T. & DUMOULIN, S. O. 2013. Macular degeneration affects eye movement behavior during visual search. *Frontiers in psychology,* 4**,** 579.

VASILYEV, A. 2020. Control of fixation duration during visual search task execution. *arXiv preprint arXiv:2004.10832*.

VAUGHAN, J. & GRAEFE, T. M. 1977. Delay of stimulus presentation after the saccade in visual search. *Perception & Psychophysics,* 22**,** 201-205.

VLASKAMP, B. N., OVER, E. A. & HOOGE, I. T. C. 2005. Saccadic search performance: the effect of element spacing. *Experimental brain research,* 167**,** 246-259.

VOGEL, E. K., MCCOLLOUGH, A. W. & MACHIZAWA, M. G. 2005. Neural measures reveal individual differences in controlling access to working memory. *Nature,* 438**,** 500-503.

WAGEMANS, J., ELDER, J. H., KUBOVY, M., PALMER, S. E., PETERSON, M. A., SINGH, M. & VON DER HEYDT, R. 2012. A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological bulletin,* 138**,** 1172.

WERTHEIMER, M. 1923. Laws of organization in perceptual forms. *A source book of Gestalt Psychology*.

WICKELGREN, W. A. 1977. Speed-accuracy tradeoff and information processing dynamics. *Acta psychologica,* 41**,** 67-85.

XU, Y. 2010. The impact of item clustering on visual search: It all depends on the nature of the visual search. *Journal of vision,* 10**,** 24-24.

ZELINSKY, G. J., RAO, R. P. N., HAYHOE, M. M. & BALLARD, D. H. 1997. Eye movements reveal the spatiotemporal dynamics of visual search. *Psychological science,* 8**,** 448-453.

ZELINSKY, G. J. & SHEINBERG, D. L. 1997. Eye movements during parallel–serial visual search. *Journal of Experimental Psychology: Human Perception and Performance,* 23**,** 244.

VIII. Supplementary

a. Additional figures

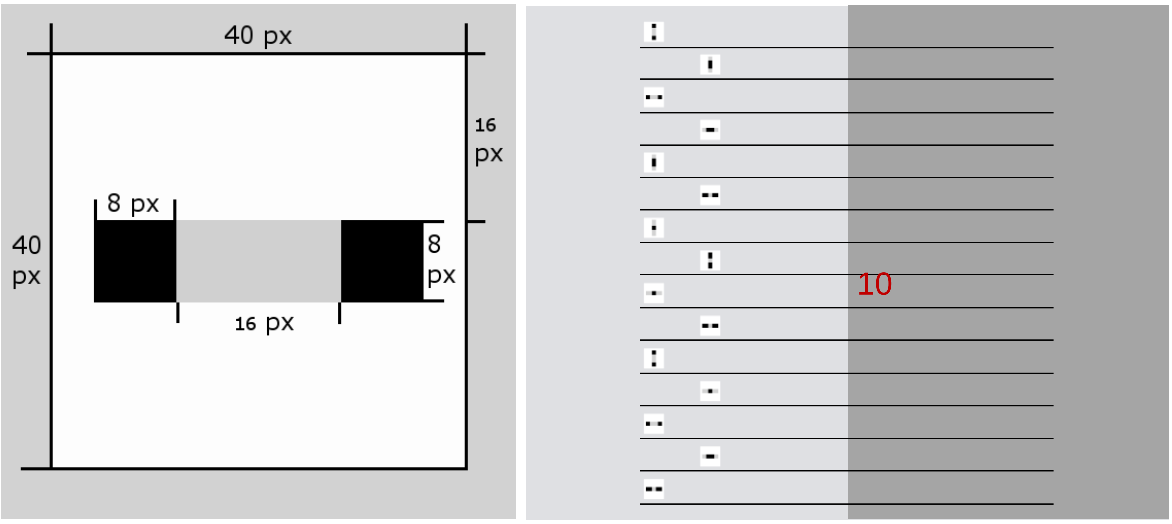

Figure 68: In detail view of a complex object and trial in Block 2
A: The complex object had the features: 1) horizontal orientation, 2) black end pieces and 3) a large gap.
B: Visualization of a trial in Block two with the number of switches left in red.

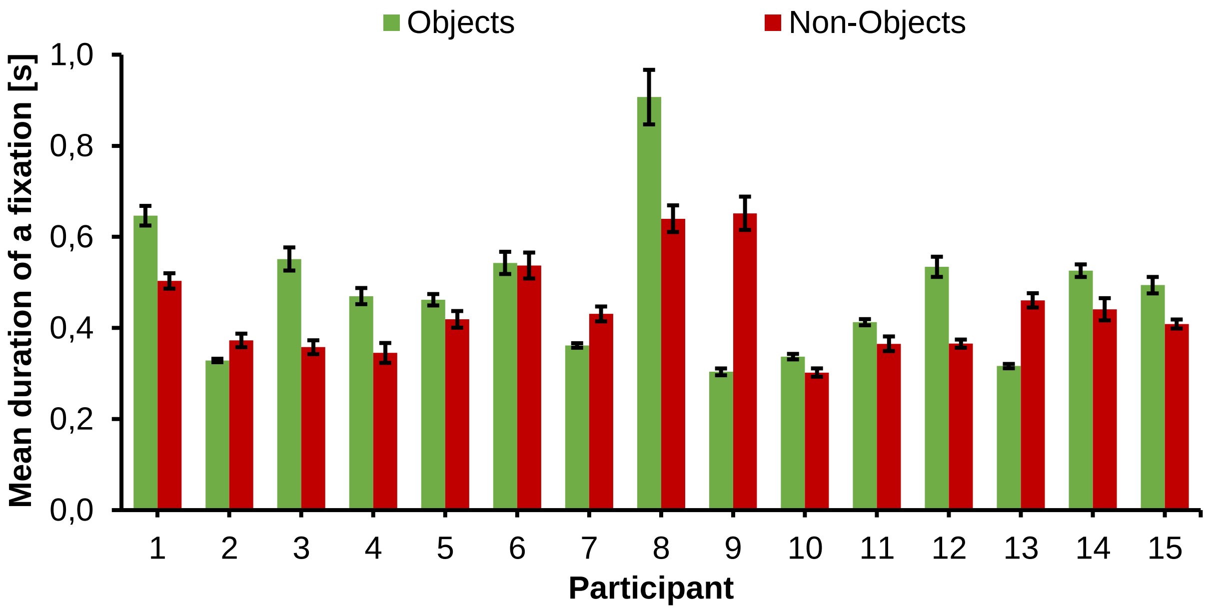

Figure 69: Mean fixation duration per participant
Mean fixation duration [s] plotted against the participant for: object (green) and non-object (red) fixations averaged over all trials (n = 40).

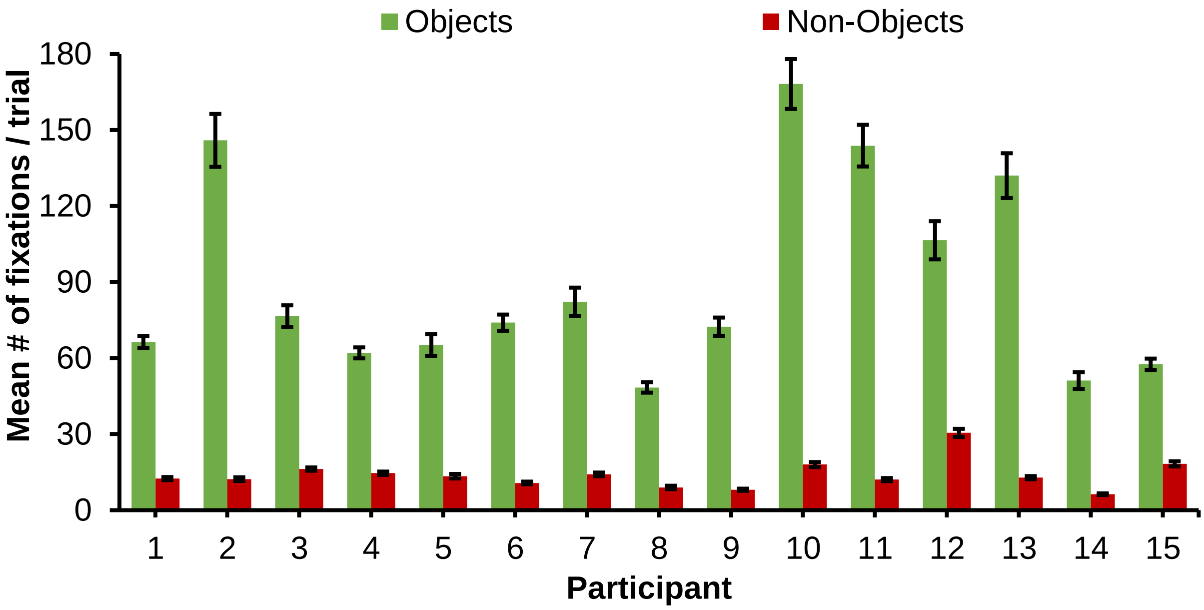

Figure 70: Mean number of fixations per participant
Mean number of fixations is plotted against the participant for: Object (green) and non-object (red) averaged over all trials (n = 40).

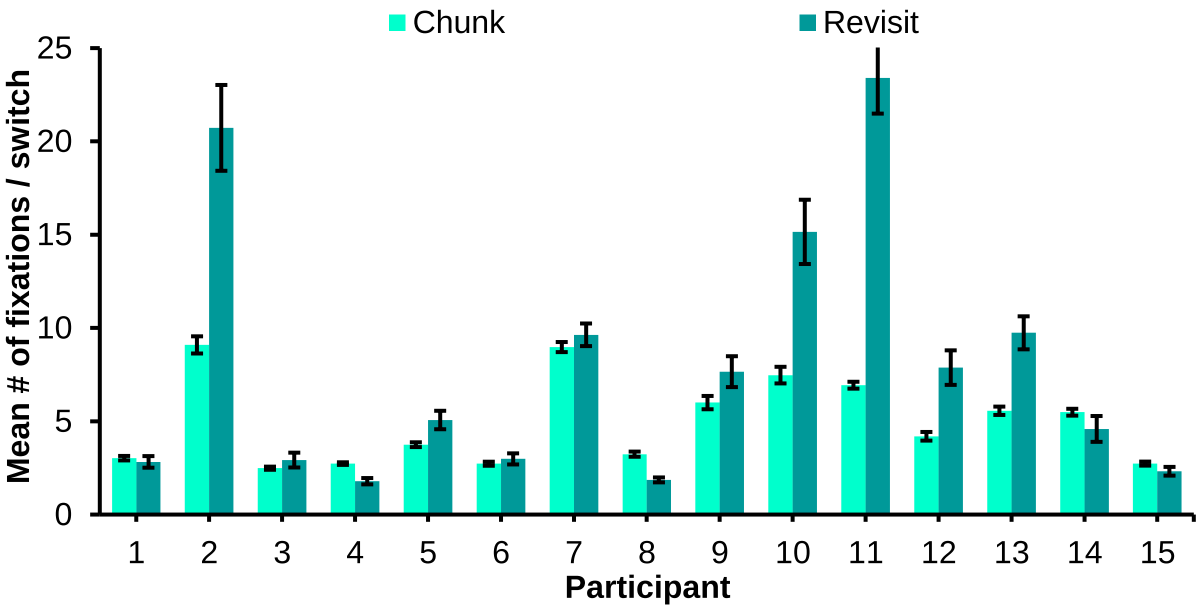

Figure 71: Mean chunk size and number of revisits
Mean number of fixations per switch is plotted against the participant for both fixation types: chunk (turquoise) and revisit (petrol) averaged over trial (n = 40).

Table 4: Tests for normal distribution of revisits depending on chunk size
The distribution of the number of revisits per chunk size was examined using Kolmogorov-Smirnov test, which was corrected using Lilliefors correction and Shapiro-Wilk test.

	Chunk size	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	Significance	Statistic	df	Significance
Revisits	1	0.395	459	< 0.001	0.650	459	< 0.001
	2	0.251	1850	< 0.001	0.752	1850	< 0.001
	3	0.223	1124	< 0.001	0.762	1124	< 0.001
	4	0.176	895	< 0.001	0.807	895	< 0.001
	5	0.187	403	< 0.001	0.790	403	< 0.001
	6	0.148	307	< 0.001	0.813	307	< 0.001
	7	0.166	183	< 0.001	0.791	183	< 0.001
	8	0.143	179	< 0.001	0.906	179	< 0.001
	9	0.138	118	< 0.001	0.836	118	< 0.001
	10	0.152	71	< 0.001	0.857	71	< 0.001
	11	0.135	64	0.005	0.914	64	< 0.001
	12	0.124	61	0.021	0.867	61	< 0.001
	13	0.170	36	0.010	0.942	36	0.060
	14	0.164	44	0.005	0.886	44	< 0.001
	15	0.099	65	0.186	0.918	65	< 0.001

b. Index of figures

c. Index of tables