

Running Head: Fixation durations as an early indirect index of concealed knowledge

Through the eyes to memory: Fixation durations as an early indirect index
of concealed knowledge

Charlotte Schwedes¹, and Dirk Wentura¹

¹Saarland University, Saarbrücken, Germany

Correspondence address:

Charlotte Schwedes
University of Saarland
Department of Psychology
Campus A2 4
66123 Saarbrücken

Phone: 0049 681 302-4641
Fax: 0049 5121 883 479
E-mail: c.schwedes@mx.uni-saarland.de

Abstract

In three experiments, we investigated an early memory effect in eye fixations, namely increased durations of the second fixations to known relative to unfamiliar stimuli. This effect occurs even if knowledge of the stimulus is deliberately concealed. In Experiment 1, we found the early memory effect using object materials and a gaze-contingent stimulus presentation that controlled for parafoveal stimulus processing. In Experiment 2a, we looked for the effect under conditions commonly used in the concealed information test (CIT). To this end, participants encoded the “to-be-concealed” knowledge incidentally while doing a mock crime task, which was followed by a CIT. Beyond the control of parafoveal stimulus processing in Experiment 1, this procedure allowed to minimize influences of carry-over processes associated with the preceding stimulus. Experiment 2b replicated Experiment 2a but applied a one-week retention interval between the encoding of the to-be-concealed knowledge and the CIT. We observed an early memory effect in all experiments, suggesting that the effect is robust, irrespective of the paradigm, stimulus materials, and retention interval used.

Keywords: eye-movements; fixation duration; memory; concealed information test;

Through the Eyes to Memory: Fixation durations as an early indirect index of concealed knowledge

There is growing interest in indirect indices of memory for previous experience. In this regard, many studies have shown that eye-movements are suitable as indirect measures of memory (for a review, see Hannula et al., 2010). Memory can be inferred, for example, from the number of fixations to certain stimuli, or from the duration of those fixations. Familiar faces (Althoff & Cohen, 1999), buildings (Althoff et al., 1999) or scenes (Ryan, Althoff, Whitlow & Cohen, 2000) typically attract fewer fixations compared to unfamiliar ones. The duration of the fixations, however, increases with previous experience. Using faces, Ryan, Hannula, and Cohen (2007) as well as Schwedes and Wentura (2012) found longer fixation durations to known compared to unknown faces. This memory effect occurred very early: Ryan et al. found the effect already in the duration of the first fixation; Schwedes and Wentura detected a small old/new effect in the duration of the first fixation and a strong effect in the duration of the second fixation. That is, an indirect index of memory can be obtained by using the duration of single fixations, within the first second of viewing a stimulus.

Compared to Ryan and colleagues (2007) who asked participants to select each known face, Schwedes and Wentura (2012) looked for the memory effect when participants tried to conceal the knowledge of a face.

Participants in this study were familiarized with two groups of faces: faces of their “foes” and faces of their “friends”. The knowledge of the friends’ faces should be concealed whenever they were confronted with them. In a following phase, participants were concurrently presented with six faces in a circular array. Five of the six faces were always unfamiliar “fillers”. The sixth face was either (a) a *target*: a known face of a “foe” that participants had to select, (b) a *probe*: a known face of a “friend” for which participants had to conceal knowledge, or (c) an *irrelevant face*: an unknown filler face.¹ In conditions (b) and (c), participants were instructed to randomly select an unknown face. A difference between the duration of fixations to “*probes*” and “*irrelevants*” should be driven by a “pure” memory effect (i.e., an effect that holds even if the intention is to conceal knowledge) as the contribution of response intention effects² should be minimized compared to a condition where the known face should also be selected.

This memory effect was found for the total fixation duration, but most importantly, it was already clear in the duration of the second fixation. It could be the case, that the memory effect in the total fixation time can be consciously avoided by intentionally directing less viewing time to a specific stimulus or by moving the eyes at a regular pace from one image to the next. Therefore, the early memory effect could be an even more

¹ To be consistent with the naming of the items in Exp. 2a and 2b reported here, we used labels that differ from the ones used in our previous study (Schwedes & Wentura, 2012).

² Here the response intention effect is defined as longer viewing times caused by the intention to select a stimulus, independent of its familiarity.

unobtrusive memory index and could be of interest for the field of indirect memory diagnostics.

The importance of the memory effect during a second fixation is based on two aspects: the time window of a second fixation and the additional information provided by the location of a second fixation. Both aspects can be related to the well-known differentiation of familiarity versus recollection-based recognition memory (for a review, see Yonelinas, 2002). According to this distinction, recognition memory can be based on the assessment of mere stimulus familiarity or on the recollection of details during the encoding episode. Note that the result found by Ryan and colleagues (2007) can be explained by mere familiarity-related memory processes since the known stimulus in a given display (if present) should always be selected. In contrast, the paradigm used by Schwedes and Wentura (2012) requires recollection-based responding as the two kinds of known stimuli, targets and probes, call for a different response. In this regard it is noteworthy that the time window of the second fixation for known faces (i.e., 266 to 678 ms post-stimulus-onset in Schwedes & Wentura, 2012) matches the typical occurrence of a recollection process, as indexed by event-related potentials (about 400-800 ms post-stimulus-onset; for a review see Rugg & Curran, 2007).

In addition, the recollection process may be supported by the additional information provided by the location of a second fixation. Hsiao

and Cottrell (2008) disentangled the confound of a longer input time and additional information provided by a second fixation location. With a gaze-contingent stimulus presentation mode, they realized a condition that allowed a stimulus presentation for the duration of two fixations with only the information of a single fixation location. Compared to the standard condition (i.e., two fixations with two different inputs) recognition performance was lower. This second input may serve as a further retrieval cue, facilitating the occurrence of recollection. Mäntylä and Holm (2006) examined the dependency of familiarity- and recollection-based recognition on the availability of different inputs. They either restricted eye-movements to one input location (approximately between the eyes) during the recognition test or allowed free viewing of the face stimuli. The restriction impaired recollection but not familiarity-based recognition. These results suggest that a second fixation, providing information from a second input location, should play a role for recollection but not familiarity-based recognition.

Based on the apparent importance of the second fixation, the early memory effect and its generalizability should be investigated in more detail. From our previous study, we established several starting points for examining the generalizability of the early memory effect.

First, the six faces in this study were always presented with a synchronous onset. Since parafoveal processing of the probe stimulus before

its first fixation was possible, the term “early memory” effect was, strictly speaking, not entirely correct. To get a more valid estimation of early effects it is important to control for parafoveal processing by gaze-contingent presentation of stimuli (i.e., a stimulus will not be shown before the gaze is directed towards its location). Using this presentation mode, we can be sure that fixation durations encompass all processing stages of a given stimulus.

Second, in our previous study we used facial stimuli. Facial stimuli are known to have special properties compared to objects in general. For example, they are processed more holistically (for an overview, see McKone & Robbins, 2011, but see Burton, Schweinberger, Jenkins, & Kaufmann, 2015). Most importantly, using faces, participants were confronted with a within-category discrimination task, a task that is more perceptually taxing than discriminating between objects of different categories. Thus, although there is no reason to assume that the basic process of discrimination by recollection is different for faces and other objects, it is conceivable that the duration of this process is shorter for objects because they are more easily discriminable. It must therefore be shown that the fixation measure we obtained is still sensitive enough to capture differences between known and unknown objects.

Third, the parallel presentation of six faces added complexities which are unnecessary given the assumed theoretical relationship between the recognition processes and the duration of the first two fixations when

looking at a stimulus. For example, there might be a gradual build-up of expectancies as the participant proceeds from (unknown) stimulus to (unknown) stimulus, knowing that the majority of trials contain a known stimulus, with unclear consequences for the assessment of fixations. Besides, there might be carry-over effects of lingering processes from one stimulus to the next one, if, for example, a stimulus holds attention (“Did I know him or not?”). With a trial by trial presentation mode, one can minimize influences on the duration of further fixations caused by enduring processes belonging to the stimulus fixated previously. The theoretical perspective of a recollection-based discrimination of targets and probes fits better with a simpler test version, which is known as the oddball paradigm.

In the oddball paradigm, participants are presented with individual stimuli trial-by-trial which have to be categorized as either belonging to a (small) target set (which had been learned before), or to a (larger) set of non-target stimuli. Probes (i.e., the to-be-concealed knowledge) are included in the non-target set. It is assessed whether probes cause different effects than irrelevant stimuli (i.e., a further subset of the unknown stimuli that serve as a control in a balanced design). Thus, it is a further desideratum to test for the early memory effect using an oddball paradigm.

Fourth, in contrast to our earlier study, a more distinct probe learning episode could increase validity. This is realised in a concealed information test (CIT; Lykken, 1959; for “oddball” versions see, e.g., Farwell &

Donchin, 1991; Seymour, Seifert, Shafto, & Mosmann, 2000): Probe knowledge is typically acquired during a mock crime scenario to make the learning context maximally different from the learning context of the targets. Thus, by increasing the ecological validity, the learning contexts for known targets and known probes are more distinct. Therefore, it might be easier to recollect the learning context or more retrieved information could be used to distinguish between targets and probes. In the study by Schwedes and Wentura (2012), participants might remember that a presented object was encoded on the computer screen, and this information could not differentiate between targets and probes, as both were learned at the screen. In Experiment 2a and 2b of the current study, however, this will be target-specific information; therefore allowing for this differentiation. Thus, in general more context specific retrieval cues are available that allow recollection-based recognition and might decrease the time needed for the effects to occur. This provides a further improvement for detecting concealed knowledge by fixation durations, especially assuming that the early memory effect is triggered by recollection-based processing.

Fifth, in the previous study we tested immediately after the learning phase for the memory effect. It is of utmost interest to test whether probe knowledge causes the memory effect even with a longer retention interval.

Overview

Experiment 1 basically used the same procedure as Schwedes and Wentura (2012). However, we used non-facial objects, and explored the early memory effect under optimized conditions using gaze-contingent stimulus presentation. This procedure eliminated parafoveal stimulus processing before the first fixation to the stimulus. Based on the results of our previous study (Schwedes & Wentura, 2012), we expected longer durations of the second fixation but not the first fixation, to probes (i.e., known objects) compared to irrelevant (i.e., unknown objects). This pattern should result in a Stimulus Type (probes vs. irrelevant) \times Fixation (first vs. second) interaction effect.

In Experiment 2a, we investigated the early memory effect in the oddball version of the CIT. Participants acquired knowledge about probes in a mock crime scenario, and were later tested with a trial-by-trial presentation of object images that had to be categorized as (rare) targets or non-targets — with the non-target list including the probes. The gaze-contingent presentation of items (as introduced in Experiment 1) was preserved in Experiment 2a. Thus, on the one hand we controlled the influence of parafoveal processing of a stimulus before it is fixated and on the other hand we minimized influences of carry-over processes associated with the preceding stimulus. Finally, to test the robustness of the early memory effect across a delay condition, Experiment 2b replicated

Experiment 2a with a one week retention interval between probe encoding and the CIT.

Experiment 1

Method

Participants. A total of thirty-six undergraduate students from Saarland University took part in the experiment in exchange for course credit. Data of two participants were excluded as they did not follow instructions (i.e., they responded incorrectly in all concealed trials). Median age of the remaining thirty-four participants (22 women, 12 men) was 23.5 years (ranging from 18 to 30 years). All had normal or corrected-to-normal vision and were native speakers of German.

Design. Experiment 1 involved displays that comprised either a target (target-display), a probe (probe-display), or an irrelevant item (irrelevant-display) intermixed with five unknown filler items. The targets served as task relevant stimuli but were not of theoretical interest. Therefore, the focus was on a 2 (stimulus type: probes vs. irrelevants) \times 2 (fixation: first vs. second) within-participants design. The assignment of objects to the three stimulus type conditions was counterbalanced across participants.

The stronger increase in the fixation duration from the first to the second fixation for probes compared to irrelevants in Schwedes and Wentura (2012) had an effect size of $d = .54$. The minimal sample size to

detect an effect of that size in the present study—with α set to .05 (two-tailed) and power set to $(1 - \beta) = .80$ —was calculated as $N = 29$ (using G*Power3; Faul, Erdfelder, Lang, & Buchner, 2007). As different stimuli were used as well as a slightly modified procedure, we decided to increase the sample size to $N = 34$, which allows detection of a medium-sized effect of $d = .50$.

Material. The stimuli comprised 108 gray-scaled images of daily objects. All objects were placed against a uniform gray background and measured 174×190 pixels. Images were organized into 18 sets of six objects. In each set, one of the six objects served as a probe, a target, or an irrelevant item and the remaining five objects served as fillers. We created three lists—A, B, and C—of each six sets for counterbalancing.

Apparatus. Eye-movements were recorded with an SMI Hi-Speed Eye-Tracker with a sample rate of 500 Hz and a spatial resolution of 0.01° . Stimuli were presented with a Windows-based computer on a 17” monitor with a resolution of 1024×768 pixels and a refresh rate of 75 Hz, using the experimental software E-Prime 2.0 Professional. The viewing distance measured 64 cm. The parameters for fixation detection were set to the default values of the eye-tracking software SMI BeGaze: the maximal dispersion value was set to 100 pixels and the minimum fixation duration to 80 ms.

Procedure. The experiment consisted of four phases: a virtual mock crime, a study phase, an experimental phase, and a follow-up test.

The virtual mock crime. In the beginning of the experiment participants had to put themselves in the position of a burglar who is breaking into an apartment. They were presented with a virtual living room on the screen that contained objects (e.g., a vase, a lamp, etc.); six of them were moveable, and participants were instructed to “steal” the six objects by dragging them into a virtual bag on the screen with the mouse. These six objects served as the probes in the later experimental phase.

The learning phase. Initially, participants were presented with an array of twelve objects—the six stolen ones (i.e., the probes) and six further objects declared as “gifts” (i.e., the targets). To familiarize participants with the objects and their probe/target status, participants then engaged in a classification task. The probes were moved to the upper right corner of the screen and the targets were moved to the upper left corner. One of the twelve objects was presented centrally on the screen, and participants had to categorize the object as either a stolen one (by clicking on a “geklaut” [“stolen”] button) or a gift (by clicking on a “geschenkt” [“received as a gift”] button). Error feedback was given in case of an incorrect categorization. All twelve objects were presented four times in random order. In the subsequent blocks of trials, this procedure was repeated without the presence of the objects in the upper corners of the display. The

task ended after the first error-free block that followed two mandatory blocks of trials without the objects in the upper corners³.

The experimental phase. In this phase the eye-movements of participants were recorded. First, the standard thirteen-point calibration procedure was administered. Then, participants were instructed to imagine that someone had given the police a tip-off accusing them of breaking into the apartment, and that they were now at a police station facing a test to find out if they had any crime-specific knowledge. Participants' task was to identify objects they had received as gifts (i.e., targets) while concealing

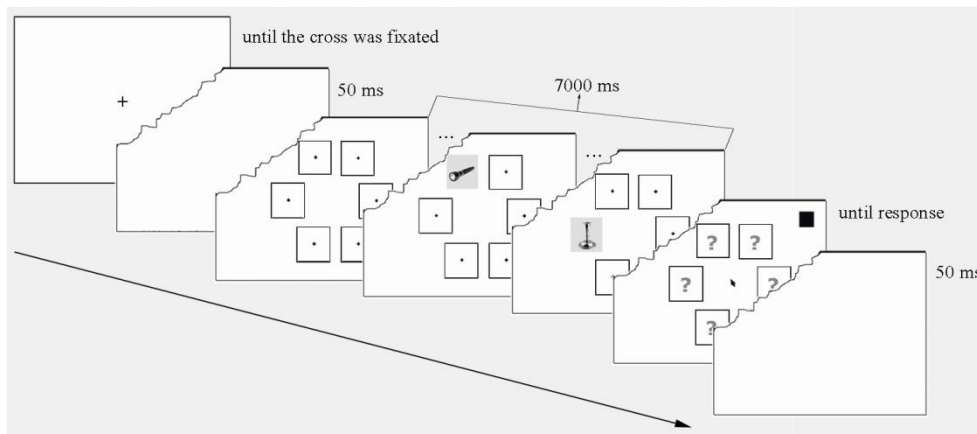


Figure 1: Example of the trial sequence in the experimental phase of Experiment 1.

knowledge about stolen items (i.e., probes). Participants were presented with circular displays containing images of six objects, five fillers and either one target (target-display), one probe (probe-display) or one irrelevant item (irrelevant-display) which, however, were masked unless fixated (see Figure

³ The number of presentations of each object was on average $M = 7.15$ ($SD = .44$). The size of the memory effect was not influenced by the amount of exposure in the learning phase.

1). Responses had to be given after the presentation of the image display, when objects were replaced with permanent masks. If there was a target present in the display, participants had to click on the location of the (now permanently masked) target. If no target was present (i.e., either a probe or an irrelevant display), participants were instructed to indicate that they did not own any of the presented objects, and click on a black button presented in the upper right corner of the screen.⁴ The participants were also told to behave inconspicuously so as not to stand out as the burglar because of their behavior. There were three practice trials to familiarize participants with the procedure of the experimental phase. Each practice trial either contained six objects or five objects and one animal; participants had to identify the animal, or click on the black button if there was no animal.

Each trial started with a central fixation cross that had to be fixated to proceed to the next display. Whenever a drift-correction was needed, it was applied at the time of the fixation cross. A 50 ms blank display followed. Then the six objects were presented, arranged in a circle. All six images were masked by white rectangles with a black frame and a small dot at the center. Participants were instructed to look at the dot to unmask the object behind the rectangle. Whenever a participant's gaze fell inside one of the frames, the mask was removed; the object was masked again when the gaze fell outside the frame. Each object was centered at the position of the

⁴ This is a change in the procedure compared to Schwedes and Wentura (2012) where participants were instructed to select one of the fillers in a probe- or irrelevant-display.

dot, as focusing the center of an object is the best position for object recognition (Foulsham & Kingstone, 2013). After seven seconds, all images were masked with permanent masks showing question marks in all frames. In addition, the mouse cursor appeared in the middle of the display and a black button appeared in the upper right corner. Participants responded by clicking on the question mark mask that covered the to-be-selected object or the black button. After the response, a blank display was presented again for 50 ms before the next trial started.

There were six target trials, six probe trials, and six irrelevant trials, presented in random order. In each display, the position of the objects was selected at random.

The follow-up test. To check participants' object knowledge, all probe and target displays were presented again without gaze-contingent presentation. Participants performed a two-stage task: they first identified the known object in each display (by clicking on the corresponding image) and then decided whether it was a stolen object or a gift. At the end of the experiment, participants filled in a questionnaire to check if they had pre-experimental knowledge of any of the specific objects and if they had tried to make use of any strategies in the experimental phase.

Data preparation. We discarded trials with incorrect responses (1.5% of all probe trials, 2.5% of all target trials, and 0.5% of all irrelevant trials) and trials with probes or targets that were not recognized in the

follow-up test (0.5% of probe trials; no target trials). In addition, we only analyzed fixation durations of images that received at least two fixations, therefore excluding 17.1% of all trials (19.0% of probe trials, 14.4% of target trials, 18.1% of irrelevant trials). As a second fixation we used both types of possible second fixations to an object: second fixations that immediately followed the first fixation to the object, as well as re-fixations (second fixations after having looked elsewhere in the display subsequently to the first fixation). We excluded outliers separately for each fixation and stimulus type according to Tukey's criterion (Tukey, 1977; i.e., values three interquartile ranges above the third quartile; 1.4% of trials). The data for the following analyses were aggregated separately for each participant and condition using the arithmetic mean. Detailed information about the range of trials carried forward for analyses in Experiment 1 are depicted in Table A2.

To gain insight into the validity of the memory effect in the second fixation to reveal someone's crime knowledge in the CIT, we include a classification analysis using the area under the receiver operating characteristic (ROC) curve. Comparable to other CIT studies (e.g., Gamer, Kosiol & Vossel, 2010; Peth, Kim & Gamer, 2013), we first converted the raw data of each participant to standardized difference scores in order to eliminate individual differences in the duration of a second fixation. For each display condition (probe-, and irrelevant-display) the mean duration of

the second fixation of the five filler stimuli was subtracted from the duration of the second fixation to the relevant probe or irrelevant stimuli. This difference score was then divided by the standard deviation of the five fillers in the corresponding display. We then aggregated these standardized differences separately for each participant and display condition. Subsequently, the scores of the probe-displays were used as values for the “guilty” condition and the scores of the irrelevant-displays for the “innocent” condition. Thus, each participant is evaluated in both conditions. These data were then used in a (ROC) analysis to estimate the area under the ROC curve (AUC) and the corresponding 95% confidence interval using the package pROC (Robin et al., 2011) in R (R Core Team, 2013).

Results

Given that only 53.4% of all images got a third fixation, we only analyzed the duration of the first and second fixation. The mean fixation durations for the first two fixations and the total fixation duration, as a function of stimulus type, can be seen in Figure 2 (see Table A3 for more detailed information).

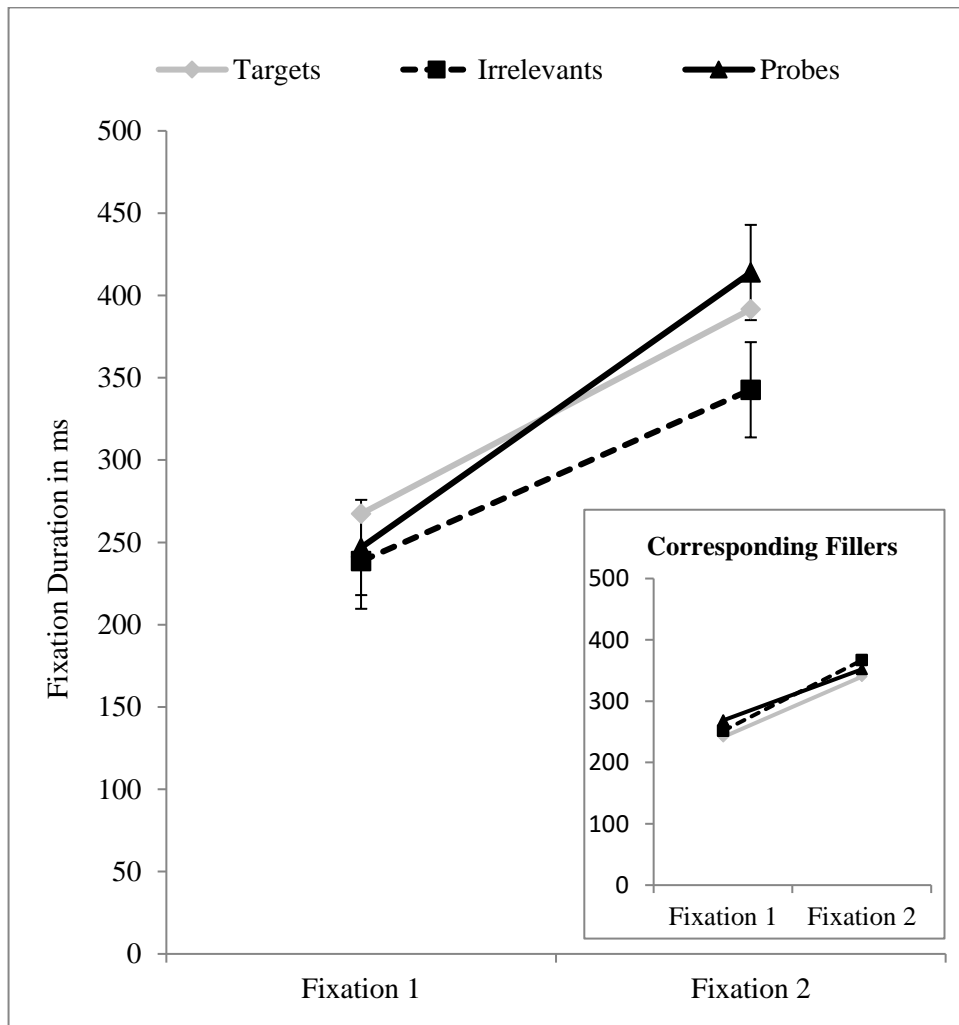


Figure 2: Fixation duration (in ms) for the first two fixations as a function of stimulus type (Experiment 1). Error bars are 95% within-subject confidence intervals (Jarmasz & Hollands, 2009) for the Fixation (first versus second) \times Stimulus Type (probes versus irrelevants) interaction. Small figure: fixation duration (in ms) for filler that were presented together with the specified stimulus types (target, probe, and irrelevant).

Probes vs. Irrelevants. We conducted a 2 (stimulus type: probes vs. irrelevants) \times 2 (fixation: first vs. second) repeated measures ANOVA on

the average duration of single fixations as the dependent variable to test our central hypothesis.⁵

Both main effects reached significance, $F(1,31) = 9.63, p = .004, \eta_p^2 = .24$ for stimulus type, and $F(1,31) = 23.69, p < .001, \eta_p^2 = .43$ for fixation; they were qualified by an interaction, $F(1,31) = 4.76, p = .037, \eta_p^2 = .13$.⁶ Planned comparisons between probe and irrelevant fixation times showed no difference for the first fixation, $t(33) = 0.58, p = .568, d = .096, 95\% \text{ CI } [-0.241, 0.457]$, but a significant difference for the second fixation, $t(33) = 3.10, p = .004, d = .529, 95\% \text{ CI } [0.182, 0.880]$. This early memory effect in the duration of the second fixation is consistent with the results of Schwedes and Wentura (2012).⁷

Targets vs. Irrelevants. We conducted a 2 (stimulus type: targets vs. irrelevant) \times 2 (fixation: first vs. second) repeated measures ANOVA to look for a memory effect for the target stimuli. Besides a main effect for fixation, $F(1,31) = 14.54, p = .001, \eta_p^2 = .32$, the main effect for stimulus

⁵ We added counterbalancing group as a between-participants control factor in order to use the correct error term (see Pollatsek & Well, 1995) and ensure that results were not affected by (slight) differences in group size (following exclusion of two participants, see *Participants* section). Results were essentially the same without adding this factor. For the sake of brevity, we will not report the results for this factor.

⁶ In the analysis of trials with only immediate second fixations to probes (86.42% of trials with a second fixation) and irrelevant (85.25%) the pattern was essentially the same, $F(1,30) = 9.00, p = .005, \eta_p^2 = .231$ for the main effect of Stimulus Type, $F(1,30) = 32.63, p < .001, \eta_p^2 = .521$ for the main effect of Fixation and $F(1,30) = 5.55, p = .025, \eta_p^2 = .156$ for the interaction effect.

⁷ The result was further corroborated by a significant three-way interaction in a 2 (display type: probe vs. irrelevant-display) \times 2 (fixation) \times 2 (stimulus type: probe/irrelevant vs. filler) repeated measures ANOVA, $F(1,31) = 5.90, p = .021, \eta_p^2 = .16$. Thus, the result does not generalize to fillers.

type was significant, too, $F(1,31) = 9.57, p = .004, \eta_p^2 = .24$. Targets were associated with longer fixations compared to irrelevant. The main effects were not qualified by a Stimulus Type \times Fixation interaction, $F(1,31) < 1$.⁸

The main effect of stimulus type in combination with the non-significant interaction indicates a tendency to a memory effect already in the duration of the first fixation. Therefore, we conducted two post-hoc t-tests (with Bonferroni-adjusted alpha = .025). A t-test for the comparison of targets versus irrelevant for the duration of the first fixation was not significant, $t(33) = 1.95, p = .060, d = .334, 95\% \text{ CI } [-0.015, 0.683]$. As target and probe stimuli were both encoded in almost the same manner during the learning phase (the only difference was that probe stimuli occurred additionally in the visual mock crime), we run a second post-hoc t-test with known stimuli (i.e., targets and probes collapsed) against unknown ones (i.e., irrelevant) to look at a potential memory effect in the duration of first fixations with more power. This test was not significant as well,, $F(1,31) = 2.22, p = .147, \eta_p^2 = .067$.

Validity of the CIT. To provide insight into the validity of the second fixation to detect concealed knowledge, we conducted ROC analyses. An AUC of .5 represents a differentiation between “innocent” and “guilty” participants at chance level, an ACU of 1 indicates perfect

⁸ The analysis of trials with only immediate second fixations to targets (80.25% of trials with a second fixation) and irrelevant (85.25%) revealed a main effect of Stimulus type, $F(1,30) = 11.30, p = .002, \eta_p^2 = .274$ as well as a main effect of Fixation, $F(1,30) = 29.30, p < .001, \eta_p^2 = .494$. The interaction effect failed to reach significance, $F(1,30) = 2.82, p = .104, \eta_p^2 = .086$.

classification. The ROC analyses revealed a differentiation between “innocence” and “guilty” above chance level when using the duration of the second fixation as a predictor, $AUC = .69$, 95% CI [.59 - .82].

Total fixation duration. For the sake of completeness, we analyzed the total fixation duration with two post-hoc t-tests (with Bonferroni-adjusted $\alpha = .025$). There was no significant difference between probes and irrelevants in the total fixation duration, $t(33) = 0.64$, $p = .527$, $d = .106$, 95% CI [-0.239, 0.458], emphasizing the importance of the early memory effect in order to detect concealed object knowledge. Targets, however, were associated with a significantly longer total fixation duration compared to irrelevants, $t(33) = 3.16$, $p = .003$, $d = .541$, 95% CI [0.192, 0.890].

Discussion

Using non-facial objects and a gaze-contingent procedure that controls for parafoveal stimulus processing, we found an early memory effect, replicating our earlier work (Schwedes & Wentura, 2012). The second fixation to a probe (i.e., a known item for which knowledge is concealed) was longer than the second fixation to an irrelevant stimulus (i.e., an unknown control item). Thus, the early memory effect in fixation durations was found even when parafoveal item processing was precluded; moreover, it does not depend on the type of the stimuli used. If this early memory effect is due to the occurrence of recognition processes during the

second fixation, it is indeed plausible that the effect is independent of the stimulus type. We will discuss this point in detail in the General Discussion.

In contrast to the early memory effect for concealed knowledge, the total fixation time did not differ between probes and irrelevants. This result is in contrast to the findings of Schwedes and Wentura (2012). The used gaze-contingent procedure with six image displays might have resulted in a regularly paced movement of the eyes from one image to the next to uncover each image at least once. This might have resulted in comparable total fixation times. A second possible explanation relates to a slight change in the procedure. Participants had to click the black button in the case of a non-target-display, whereas in our previous experiment a filler had to be arbitrarily selected in these trials. Thus, in the latter case, participants had to remember the probes' position to not accidentally select the probe; this additional encoding might have influenced later fixations.

In addition, receiver operating analyses revealed a differentiation between “innocence” and “guilty” above chance level when using the duration of the second fixation as a predictor.

The analysis for to-be-revealed knowledge (known and selected stimulus in target-displays) showed longer durations to known targets compared to irrelevants within the first two fixations. This indicates already a tendency toward a memory effect in the duration of the first fixation. However, a supplementary analysis with known stimuli (targets and probes

collapsed) against unknown ones (irrelevants) showed no memory effect in the duration of first fixations to known stimuli. Thus, robust early memory effects did not appear before the duration of second fixations.

The results of Experiment 1 suggest that the duration of the second fixation can reveal someone's object knowledge. In Experiments 2a and 2b, we investigated the early memory effect in an oddball paradigm with a distinct probe learning episode. That is, we adapted the concealed information test (CIT) to our experiments. As far as we are aware, there has only been one study that has used fixation durations in a CIT with a mock crime scenario (Peth, Kim, & Gamer, 2013). These authors only analyzed the total fixation times and replicated the memory effect. However, stimulus presentation conditions in that study did not allow for an analysis of the duration of single fixations, and thus of the early memory effect.

Experiment 2a and 2b

In Experiment 2a and 2b, we used the “oddball” variant of the CIT. After committing a mock crime, participants were confronted with three different kinds of stimuli: designated targets (not crime-related) and non-targets that are either crime-related (probes) or not crime-related (irrelevants). On each trial, participants are presented with one of these stimuli and have to classify it as a target or a non-target. For someone without any knowledge of the mock-crime, this task is a simple discrimination task between known (targets) and unknown stimuli. For

“guilty” participants, however, it is a more difficult task because they know both the targets and the crime-related probes, and have to discriminate within the set of known stimuli to make the correct categorization. This “oddball” variant is often used in reaction time-based CITs (e.g., Seymour & Kerlin, 2008; Seymour, Seifert, Shafto, & Mosmann, 2000; Verschuere, Crombez, Degrootte, & Rosseel, 2010) and P300-based CITs (i.e., CITs that focus on a specific ERP component that is evoked by meaningful or rarely presented items about 300 ms after stimulus onset; e.g., Farwell & Donchin, 1991; Rosenfeld et al., 1988; Rosenfeld, Rao, Soskins, & Miller, 2002) and has unveiled different patterns between probes and irrelevant.

Using a classical oddball CIT with its trial-by-trial item presentation, we took the advantage of minimizing the complexity of the display and avoided possible influences on the fixation duration that do not belong to the processing of the fixated stimulus (i.e., carry-over effects).

The different learning contexts of the probes (during a mock crime) and the targets (during a study phase) did not only enhance the ecological validity of the test. They also make it easier (and thereby faster) to recollect the item specific context that is necessary to make a source discrimination between familiar targets and familiar probes in order to respond in the correct way. Thus, this setting poses a further improvement to a CIT based on the duration of early fixations.

In Experiment 2a, the CIT followed directly after participants acquired the crime-related knowledge by committing a mock crime. We still expected an early memory effect in the duration of the second fixation for to-be-concealed knowledge. That is, we expected a stronger increase in fixation duration from the first to the second fixation for known probes compared to unknown irrelevant objects.

In Experiment 2b, we administered the CIT one week after the participant committed the mock crime, with everything else being equal to Experiment 2a. As the effect in the duration of fixations is regarded as a long-term memory effect, it is important to test the robustness of the early memory effect with a longer retention interval.

For exploratory reasons we measured participants' arousal level to investigate whether the early memory effect is moderated by the arousal induced by the mock crime or the CIT.

Method

Participants. Two different samples each comprising of forty undergraduate students from Saarland University took part in the two experiments. In Experiment 2a, we had to exclude three participants; two participants showed “staring” behavior—with second fixations that lasted longer than the stimulus presentation in most of the trials, and one further participant responded incorrectly in 49% of the trials. The median age of the remaining thirty-seven participants in Experiment 2a (23 women, 14 men)

was 23 years (ranging from 19 to 33 years) and in Experiment 2b (27 women, 13 men) the median age was 22 years (ranging from 18 to 28 years). All participants had normal or corrected-to-normal vision and were native speakers of German. All participants received a monetary compensation of €8 for their participation.

Design. Both experiments employed a 2 (stimulus type: probe vs. irrelevant) \times 2 (fixation: first vs. second) within-participants design. We prepared two different mock crime scenarios in order to counterbalance the materials between the probe and irrelevant trials (see *Materials*).

The stronger increase in the fixation duration from the first to the second fixation for probes compared to irrelevants in Experiment 1 had an effect size of $d = 0.46$. To detect such an effect with $\alpha = .05$ (two-tailed) and a power of $1 - \beta = .80$, we calculated a required sample size of $N = 40$ for each Experiment (2a and 2b), using G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007).⁹

Material and Apparatus. In both experiments, we used exactly the same 112 colored images of objects. All were placed against a uniform gray background and measured 270×270 pixels. Ninety of these objects (15 targets, 15 probes, 15 irrelevants, and 45 fillers) were used in the CIT. The remaining stimuli served as distractors in the target-learning phase of the CIT (16 objects), or as objects in the practice trials of the CIT (one target,

⁹ As noted in the Participants section, we had to discard three participants in Experiment 2a. Actual power was therefore $1 - \beta = .78$.

one probe, one irrelevant, and three fillers). We created two lists—A and B—of 15 objects each for counterbalancing materials between probes and irrelevant. Eye-movement recording settings were identical to Experiment 1, except that stimuli were now presented on a 24” monitor with a refresh rate of 100 Hz and a resolution of 1920 × 1080 pixels.

Procedure. The experiments comprised three phases: a mock crime, the CIT, and a follow-up test.

The mock crime. After arriving in the lab and signing a consent form, participants rated their current arousal level by filling in the arousal subscale of the Self-Assessment Manikin (SAM; Bradley & Lang, 1994). This first assessment served as a baseline.

Then participants were told that they would now have to go to one of two rooms because Experimenter 1 had bet Experimenter 2 €10 that she would be able to detect which room the participant had been in by using the gaze behavior in a later task. Participants were instructed to choose one of two envelopes, in which they found a room number and a crime task list that they would have to execute in the room (e.g., to change research results). They were told that the two experimenters did not know which room they had drawn. In contrast to Experiment 1 but in line with studies using the CIT, we provided an incentive for the participant to conceal the drawn room number and their knowledge of the mock crime in the following CIT.

Experimenter 2 offered to share the €10 with them if Experimenter 1 failed to guess the room correctly.

If the participant had no further questions, they were asked to leave the lab, read the instructions they found inside the envelope, and go to the designated room. They were instructed to knock on the door to ensure there was no one inside. In the room, they had to execute the mock crime by completing their task list, thereby interacting with a set of objects (see Table A1; for example, they had to open the pencil case/document wallet to take a letter out of it or they used a permanent marker/white-out to obscure the telephone number on the letter). They had to fill in a second SAM, measuring their arousal during the mock crime. Immediately after the mock crime, they had to return to the lab.

In Experiment 2a, Experimenter 1 told the participant that she would now try to find out which room they had been in by conducting a CIT. In Experiment 2b, the participants were told that the first session was over and that they had to return to the lab in a week's time to take part in the CIT. When they returned after a week, they first filled in an additional SAM to measure their baseline arousal in the second session, before the CIT was conducted.

The CIT. In the second part of the experiment, the participants first had to learn the target objects of the CIT. In this study phase, they were

presented with 16 target objects¹⁰ (concurrently) on the computer screen and had to memorize them. To ensure that the targets had been encoded properly, participants then performed a categorization task, which intermixed the 16 targets with 16 new objects. Items were presented centrally one-by-one, and participants had to categorize them as targets (by pressing the X-key) or non-targets (M-key). Error feedback was given in case of incorrect categorization. A block of trials comprised all 32 objects. The procedure stopped after the first error-free block that followed two mandatory blocks of trials.¹¹

In the main phase of the CIT, the eye-movements of participants were recorded (after the standard calibration procedure). Each trial in the CIT started with a central fixation cross that had to be fixated to proceed to the next display. After a 50 ms blank display, a frame with a dot inside was presented to the left or right of the fixation cross. The participants were instructed to move their gaze to the dot inside the frame. When their gaze position crossed the border of the frame, the image of an object was presented in the frame, with the object centered on the position of the (no longer visible) dot. The participants were instructed to look at the object as long as it was presented. After three seconds, a display prompted participants to give their target (X-key) or non-target (M-key) response.

¹⁰ Note, 15 target objects were used for the main phase of the CIT; one target was used for practicing the CIT; see above and below.

¹¹ Each target was presented on average $M = 3.2$ ($SD = .43$) times in Exp. 2a and $M = 3.2$ ($SD = .51$) times in Exp. 2b. Adding target exposure as a covariate in the analyses reported below did not change the memory effect.

After the response, a blank display was presented again for 50 ms before the next trial started (see Figure 3). Six practice trials were given to familiarize participants with the procedure of the CIT.

In the main phase, there were fifteen trials (1/6 of all trials) each with a target, a probe, and an irrelevant object, and 45 trials (3/6 of all trials) with a filler object, presented in randomized order. In each trial, the position of the object (left versus right) was selected at random.

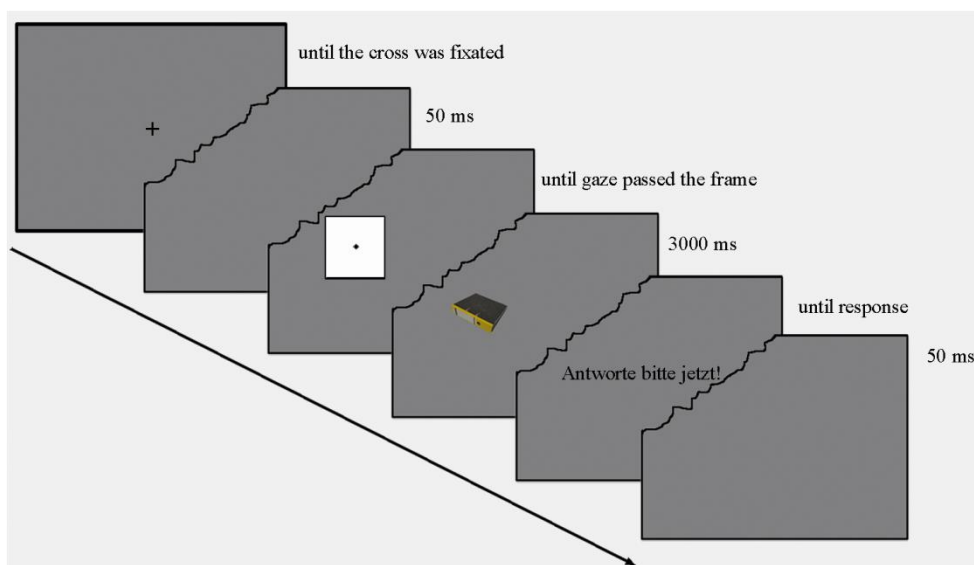


Figure 3: Example of the trial sequence in the experimental phase of Experiments 2a and 2b.

The follow-up test. Subsequently, we checked participants' knowledge of the probe objects. To this end, they were presented with 15 trials, each containing a circular display with 6 items (one target, one probe, one irrelevant, and three fillers). The participants' task was to correctly identify the object they knew from their mock crime (i.e., the probe) and to

select it with a mouse click. Finally, participants had to indicate how aroused they were during the CIT by filling in a final SAM.

At the end of the experiment, participants filled in a questionnaire to check if they had pre-experimental knowledge of any of the objects and if they had tried to make use of any strategies in the CIT.

Data preparation. Trials of the CIT were excluded from analysis if a participant had made a mistake in their target/non-target categorization (Exp. 2a: 1.0%, Exp. 2b: 0.9% of all experimental trials) or if they had responded incorrectly in the follow-up test (Exp. 2a: 0.2%, Exp. 2b: 0.8% of all trials). It was not necessary to exclude items because of pre-experimental knowledge. Although most of the participants had seen similar items before (e.g., everyone had used a glue stick before), the specific items used in the mock crime were unknown to them. We only analyzed trials with at least two fixations; the percentage of images that received a first, second, or third fixation can be seen in Table A4 in the Appendix. These criteria led to the exclusion of a total of 8.8% (Exp. 2a) and 8.6% (Exp. 2b) of all trials. We excluded outliers separately for each fixation and stimulus type according to Tukey's criterion (Tukey, 1977; i.e., values that were three interquartile ranges above the third quartile). This led to the exclusion of 2.4% and 2.0% of first fixations, as well as 1.1% and 1.6% of second fixations in Exp. 2a and Exp. 2b, respectively. The data were aggregated using the arithmetic mean.

As in Experiment 1 we prepared the data of Experiment 2a and 2b for ROC analyses. We subtracted the average duration of the second fixation of all filler stimuli from the duration of the second fixation to each probe as well as irrelevant stimulus. These difference scores were then divided by the standard deviation of the duration of the second fixation of the filler stimuli. Again, each participant is evaluated in both conditions, as “innocent” with the irrelevant trials and as “guilty” with the probe trials. These data were then used in a ROC analysis to estimate the area under the ROC curve (AUC) and the corresponding 95% confidence interval using again the package pROC (Robin et al., 2011) in R (R Core Team, 2013).

Results

Probes vs. Irrelevants. The mean fixation durations are depicted in Figure 4. To analyze if the early memory effect for concealed knowledge is different after a delay, we combined the data sets of Exp. 2a and Exp. 2b. We conducted a 2 (stimulus type: probes vs. irrelevants) \times 2 (fixation: first vs. second) \times 2 (delay: immediate [Exp. 2a] vs. one week [Exp. 2a])

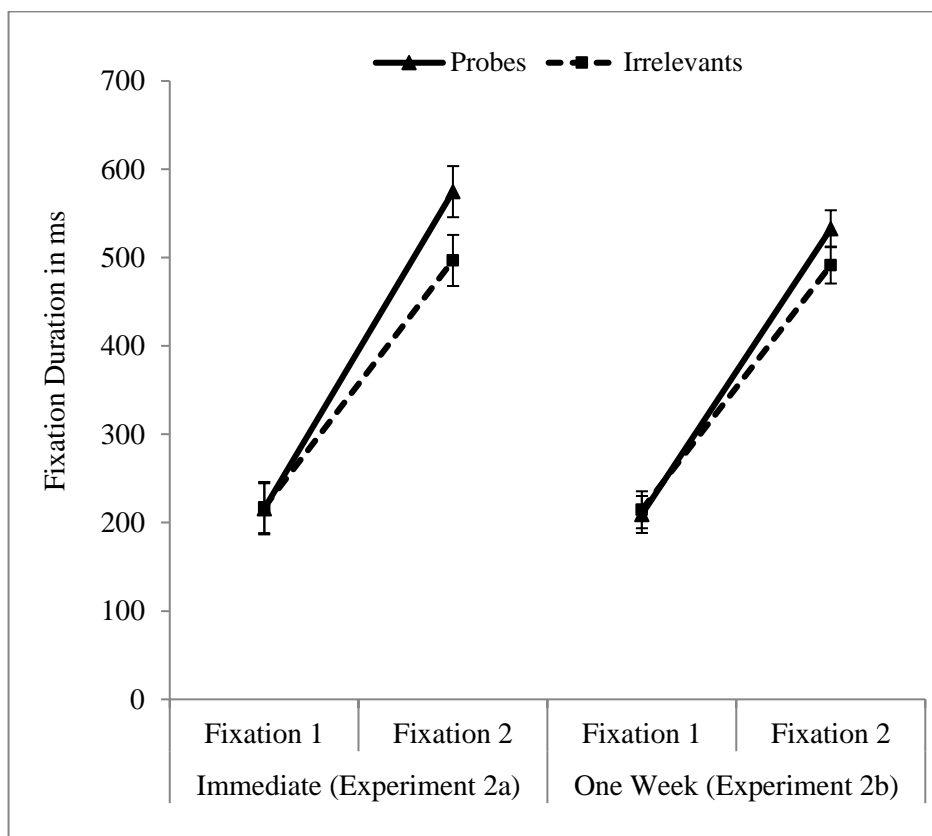


Figure 4: Left: Fixation duration (in ms) for the first three fixations, separately for probes and irrelevants in Experiment 2a. Right: Fixation duration (in ms) for the first three fixations, separately for probes and irrelevants, in Experiment 2b. Error bars are 95% within-subject confidence intervals for the interaction effect of Stimulus Type (probes, irrelevant) \times Fixation (first, second; Jarmasz & Hollands, 2009).

repeated measures ANOVA¹² with delay as a between-participant factor, and the average duration of single fixations as the dependent variable. The analysis revealed significant main effects of stimulus type, $F(1,73) = 6.66$, $p = .012$, $\eta_p^2 = .084$, and fixation, $F(1,72) = 416.42$, $p < .001$, $\eta_p^2 = .851$. These effects were qualified by a Stimulus Type \times Fixation interaction, $F(1,73) = 12.50$, $p < .001$, $\eta_p^2 = .146$. All effects involving the delay factor

¹² We again added counterbalancing group as a between-participants control factor (see Experiment 1). Results were essentially the same without adding this factor.

failed the criterion of significance, all $F_s < 1.52$, $p_s > .221$. There were no significant differences between probes and irrelevant for the duration of the first fixation, $t(76) = 0.72$, $p = .471$, $d = -.083$, 95% CI [-0.309, 0.144]. However, as predicted, the second fixation was significantly longer for probes compared to irrelevant, $t(76) = 3.13$, $p = .002$, $d = .357$, 95% CI [0.130, 0.584]. These results are in line with our hypothesis of a memory effect in the duration of the second fixation that is not significantly affected by the retention interval, as we failed to find a moderation of the memory effect by delay.

Despite the absence of any moderation by the retention interval, ultimately Experiment 2a and 2b were separate Experiments that tested for an immediate and delayed memory effect, respectively. Therefore, we ran the 2 (stimulus type: probes vs. irrelevant) \times 2 (fixation: first vs. second) ANOVA separately for Experiments. We think that separate analysis are important, as the experiments are the first ones that investigated the early memory effect in the duration of second fixations in an immediate as well as a delayed “oddball” CIT.

For Experiment 2a, the analysis revealed significant main effects of stimulus type, $F(1,35) = 4.99$, $p = .032$, $\eta_p^2 = .125$, and fixation, $F(1,35) = 171.51$, $p < .001$, $\eta_p^2 = .831$, as well as a significant interaction, $F(1,35) = 7.21$, $p = .011$, $\eta_p^2 = .171$. There was no difference between probes and irrelevant in the first fixation, $t(36) = -0.25$, $p = .806$, $d = -.041$, 95% CI [-

0.374, 0.293], but a significant longer second fixation for the probes compared to irrelevants, $t(36) = 2.58$, $p = .014$, $d = .434$, 95% CI [0.090, 0.757].

For Experiment 2b, the analysis revealed a significant main effect fixation, $F(1,38) = 260.70$, $p < .001$, $\eta_p^2 = .873$, as well as a significant interaction, $F(1,38) = 5.06$, $p = .030$, $\eta_p^2 = .118$ ($F(1,38) = 1.75$, $p = .194$, $\eta_p^2 = .044$, for the main effect of stimulus type.). Again, there was no difference between probes and irrelevants in the first fixation, $t(39) = -.72$, $p = .478$, $d = -.113$, 95% CI [-0.433, 0.207], but still a longer second fixation for probes compared to irrelevants, $t(39) = 1.79$, $p = .040$ (one-tailed), $d = .283$, 95% CI [-0.036, 0.603].

For the sake of completeness, we analyzed the total viewing time in a 2 (stimulus type: probes vs. irrelevants) \times 2 (delay) repeated measures ANOVA¹³ with delay as a between-participant factor and the total fixation time as the dependent variable. The main effect of stimulus type, $F(1,73) = 2.37$, $p = .128$, $\eta_p^2 = .031$, as well as all the other effects (F s < 1) did not reach significance.

In contrast to Experiment 1, we do not report the analyses for the targets for Experiments 2a and 2b, as the stimuli used were only counterbalanced across the probes and irrelevants. Interested readers can

¹³ We again added counterbalancing group as a between-participants control factor. Results were essentially the same without adding this factor.

find the means and standard deviations for all stimulus types separately for the first three fixations as well as for the total fixation duration in Table A4.

The analysis of the arousal level revealed a significant increase in arousal during the mock crime and the CIT compared to the baseline rating, $t(75) = 6.93, p < .001, d = .795, 95\% \text{ CI} [.567, 1.024]$, and $t(75) = 3.84, p < .001, d = .443, 95\% \text{ CI} [.211, .668]$. Since none of the results reported above concerning the memory effect were moderated by the increase in arousal, we will not report these analysis.

Validity of the CIT. Again we conducted ROC analyses to gain insight into the validity of the second fixation to detect concealed knowledge. The ROC analyses revealed an AUC of .61, 95 % CI [.53 - .70] which is significantly above .5. That is, it was possible to predict “innocence” and “guilt” using the standardized duration of the second fixation. Although there was no significant difference between the two AUCs for Experiment 2a and 2b according to a bootstrap test for unpaired samples (Robin et al., 2011), $D = 0.746, p = .457$, for the sake of completeness we report the results for both experiments separately. For Experiment 2a, the AUC was .65, 95 % CI [.53 - .76] which is significantly above .5; for Experiment 2b, the AUC was .58, 95 % CI [.46 - .70] which is not significantly above .5.

Discussion

Experiments 2a and 2b revealed an early memory effect in fixation durations under conditions common in the field of indirect memory diagnostics. The trial-by-trial presentation had the advantage of minimizing the complexity of the display. Compared to Experiment 1, it prevented possible carry-over effects from preceding stimuli. In addition, we used a mock crime scenario, which implemented a distinct encoding context for the to-be-concealed probe items, and we employed an “oddball” CIT with its clear target versus non-target categorization task. Experiment 2b investigated the early memory effect after incidental encoding and a one-week delay.

These aspects of the current paradigm contribute to the validity of the approach in that they promote a clear distinction between targets (i.e., stimuli that are only task-relevant within the CIT) and probes (i.e., the to-be-concealed knowledge). This prevented any confusion of targets and probes that might have arisen in former studies due to similarities of the learning procedures for “friends” (i.e., to-be-concealed face probes) and “foes” (i.e., to-be-revealed face targets; Schwedes & Wentura, 2012) or “stolen objects” and “gifts” (Experiment 1).

Experiment 2 replicated the early memory effect found in Experiment 1, that is, a stronger increase in fixation duration from first to second fixation for the crime-related probe stimuli compared to irrelevant stimuli. When using the duration of the second fixation for classification

analyses, a differentiation between “innocence” and “guilt” was above chance level. The memory effect was lacking when investigating the total fixation duration. This is not surprising, as in Experiment 2a and 2b, participants were instructed to look at the object as long as it was presented. Therefore, irrespective of item type, they fixated on objects approximately for the same amount of time in total. In addition, participants were given a monetary incentive to conceal their knowledge and therefore they were more motivated to try to beat the test. It is easier to influence the total viewing time deliberately than the duration of a single early fixation. The motivation to conceal knowledge may have weakened the effect in the total fixation duration, but we think the single stimulus presentation with the instruction to look at it as long as it is present is the main reason for the absent memory effect in the total fixation duration in Exp. 2a and 2b.

Our results corroborate the robustness and importance of the early memory effect even under conditions that are common in a standard CIT setting, i.e. when the crime-related stimuli are encoded only incidentally while committing a mock crime.

The analyses across Exp. 2a and 2b failed to reveal any significant moderation of the early memory effect by delay. We have to concede that the separate analysis of Experiment 2b (i.e., a CIT after a one week delay) showed only a small memory effect which, as a consequence, leads to a non-significant classification of participants.

General Discussion

The goal of the present study was to examine the early memory effect in the duration of gaze fixations on “to-be-concealed” knowledge. In contrast to previous concealed information tests (CITs) with fixation duration as the dependent variable, we used non-face objects and conditions that correspond to a standard oddball CIT procedure. By using gaze-contingent stimulus presentation (in all presented experiments) and a trial-by-trial stimulus presentation (Experiment 2a and 2b), we tried to restrict the ongoing processes in a fixation to the processes that belong to the fixated stimulus. These improvements make it easier to interpret effects in the fixation durations with ongoing processing and recognition of the fixated stimulus.

In Experiment 1, we found an early memory effect in the second fixation duration, consistent with findings of our previous study (Schwedes & Wentura, 2012), but using object stimuli and gaze-contingent stimulus presentation. The occurrence of the early memory effect in Experiment 1 speaks for the effect’s generalizability across materials. The gaze-contingent presentation mode rules out parafoveal processing, and thus allows for a more valid estimation of early effects as the first processing of the stimulus took place with the onset of the first fixation to the stimulus. The memory effect in the second fixation becomes more important as it seems possible to

conceal the “crime” knowledge in the total viewing time. There was no difference between probe and irrelevant items concerning the total fixation duration.

In a second step, we tested the memory effect under conditions usually used to detect concealed knowledge. To this end, we used the “oddball” variant of the CIT with a mock crime scenario (Experiments 2a and 2b). This allowed us to test the robustness of the memory effect for stimuli that are incidentally encoded in an arousing and more realistic situation that is clearly differentiated from the target encoding necessary for the CIT. Furthermore, with the gaze-contingent and trial-by-trial stimulus presentation we controlled the influence of parafoveal processing of a stimulus before it is fixated and we minimized influences of carry-over processes. Under these conditions, we again found the early memory effect in the duration of the second fixation. Even with a one-week delay between the encoding of the crime-related probe stimuli and the CIT (Experiment 2b), we found a weaker but still (one-tailed) significant effect. As in Experiment 1 there was no memory effect in the total fixation duration. These results support the assumption that under some circumstances the use of the early memory effect within the first two fixations to detect concealed knowledge is the more valid measure to detect concealed knowledge compared to the total viewing time, be it for procedural reasons (e.g., trial-

by-trial stimulus presentation) or for the possibility to control the total viewing time to a stimulus in a stronger fashion.

There is consensus that the duration of a fixation is associated with the duration of cognitive processes concerned with the input from the fixation (e.g., Irwin, 2004). Unema, Pannasch, Joos and Velichkovsky (2005) assumed that after an early orientation phase the inspection of informative details in a scene takes place. In the context of the present study, one can hypothesize that the processing of detailed stimulus information necessary to identify stimuli and to discriminate between different classes of known stimuli does not take place before the second fixation.

Corroborative evidence comes from research on the processing stages of object recognition, using event-related potentials (ERPs). The early pictorial and structural encoding of an object takes place within the first 150 ms post stimulus onset (see, e.g., Johnson & Olshausen, 2003; Rousselet, Husk, Bennett & Sekuler, 2008). Thereafter, the visual input is matched with stored memory representations. This process results in a familiarity signal that allows the discrimination between known and unknown stimuli, followed by a slower recollection process (see Rugg & Curran, 2007, for review). The latter process enables the retrieval of contextual details of a prior episode.

Specifically for object recognition, Miyakoshi, Nomura, and Ohira (2007), used an oddball paradigm with unfamiliar and two types of familiar objects, self-relevant and “simple” familiar (i.e., not self-relevant) objects. These stimulus types are comparable to the kinds of familiar objects used in our Experiments 2a and 2b. They also reported a first differentiation between familiar and unfamiliar objects about 200 to 300 ms after object onset. In a later time window, between 300 and 700 ms post-stimulus onset, they showed a differentiation between simple familiar and self-relevant familiar objects.

Therefore, we can assume a differentiation between targets and probes in our experiments from 300 ms after stimulus onset. The component associated with the recollection-based identification process of familiar objects reported by Miyakoshi et al. (2007) matches the time window of the second fixation, which ranges, on average, from 231 ms to 683 ms (Exp. 1: 250 ms to 631 ms; Exp. 2a and 2b, on average: 212 ms to 752 ms) post stimulus onset. We therefore assume that the ongoing recognition processes cause longer second fixations to known objects, as a recollection-based recognition is necessary to prevent errors (i.e., to prevent revealing knowledge of a “stolen” (probe) object).

Given that in our experiments participants had to distinguish between the two types of known stimuli to respond correctly and that this recollective experience might be accompanied by familiarity, the studies

cannot clarify whether a familiarity signal may suffice to provoke longer fixations, or whether the retrieval of contextual details is the causal process. To match the underlining memory processes (via their associated brain potentials) with the chronological sequence of the first and second fixations, further research is necessary that combines both methods in one study.

One aspect of our experiments, which is ignored by the time-based interpretation above, is the fact that each fixation is associated with a new input. Thus, it is possible that a second fixation has a functional significance. Hsiao and Cottrell (2008) showed in a face recognition experiment that recognition performance increased with an additional information provided by the location of a second fixation compared to a control condition with the same input duration but constant input information (the stimulus was relocated during the saccade between first and second fixation). In an unpublished experiment from our lab (Schwedes & Wentura, 2016) we used the technique of Hsiao and Cottrell (2008) and looked at the impact of the input duration as well as the availability of an additional input on the underlying recognition processes (familiarity and recollection) when recognizing faces. Both components had a significant impact on recollection-based recognition. These results underpin the functional role of a second fixation for recollection-based memory effects.

In summary, the second fixation to a familiar stimulus has an important role for memory retrieval, regardless of whether it typically

occurs in a time window that is associated with a specific processing stage during object recognition, or whether it provides the memory system with a second retrieval cue. Therefore, it makes sense that the early memory effect investigated here seems to occur predominantly in the second fixation. Our results demonstrate the robustness of the early memory effect across different paradigms and stimulus materials. After a one week retention interval the effect remains, but is somewhat reduced. Although we instructed our participants to show no difference in viewing behavior for the different kinds of stimuli, the early memory effect for to-be-concealed knowledge appeared. This emphasizes again the importance of the early memory effect and its potential as an indirect index of memory in further experiments to detect concealed knowledge. The studies reported here were not conducted to investigate if the early memory effect could intentionally be inhibited by the use of countermeasures. Further studies should be conducted to investigate if the fixation duration based CIT could be faked when participants are explicitly instructed to use countermeasures. Nevertheless, our series of experiments highlight the potential of the duration of early fixations as an indirect memory indicator.

References

- Althoff, R. R., Cohen, N. J., McConkie, G., Wasserman, S., Maciukenas, M., Azen, R., & Romine, L. (1999). Eye-movement based memory assessment. In W. Becker, H. Deubel & T. Mergner (Eds.), *Current Oculomotor Research: Physiological and Psychological Aspects* (pp. 293-302). New York: Plenum Press.
- Althoff, R. R., & Cohen, N. J. (1999). Eye-movement-based memory effect: A reprocessing effect in face perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 997-1010.
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*, 49-59.
- Burton, A. M., Schweinberger, S. R., Jenkins, R., & Kaufmann, J. M. (2015). Arguments against a configural processing account of familiar face recognition. *Perspectives on Psychological Science*, *10*, 482-496.
- Farwell, L. A., & Donchin, E. (1991). The truth will out: Interrogative polygraphy ("lie detection") with event-related brain potentials. *Psychophysiology*, *28*, 531-547.
- Faul, F., Erdfelder, E., Lang, A.-G, & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

- Foulsham, T., & Kingstone, A. (2013). Optimal and preferred eye landing positions in objects and scenes. *The Quarterly Journal of Experimental Psychology*, *66*, 1707-1728.
- Gamer, M., Kosiol, D., and Vossel, G. (2010). Strength of memory encoding affects physiological responses in the Guilty Actions Test. *Biological Psychology*. *83*, 101–107.
- Hannula, D. E., Althoff, R. R., Warren, D. E., Riggs, L., Cohen, N. J., & Ryan, J. D. (2010). Worth a glance: using eye movements to investigate the cognitive neuroscience of memory. *Frontiers in Human Neuroscience*, *4*, 1-16.
- Hsiao, J. H., & Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological Science*, *19*, 998-1006.
- Irwin, D. E. (2004). Fixation location and fixation duration as indices of cognitive processing. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 105-133). New York: Psychology Press.
- Johnson, J. S., & Olshausen, B. A. (2003). Timecourse of neural signatures of object recognition. *Journal of Vision*, *3*, 499-512.
- Lykken, D. T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology*, *43*, 385-388.
- Mäntylä, T., & Holm, L. (2006). Gaze control and recollective experience in face recognition. *Visual Cognition*, *14*, 365-386.

- McKone, E., & Robbins, R. (2011). Are faces special? In A. J. Calder, G. Rhodes, M. H. Johnson & J. V. Haxby (Eds.), *The Oxford Handbook of Face Perception* (pp. 149-176). Oxford: University Press.
- Miyakoshi, M., Nomura, M., & Ohira, H. (2007). An ERP study on self-relevant object recognition. *Brain and Cognition*, *63*, 182-189.
- Peth, J., Kim, J. S. C., & Gamer, M. (2013). Fixations and eye-blinks allow for detecting concealed crime related memories. *International Journal of Psychophysiology*, *88*, 96-103.
- Pollatsek, A., & Well, A.D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 785-794.
- R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, Austria.
Retrieved from <http://www.R-project.org/>
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C. & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*, 1-8. DOI: 10.1186/1471-2105-12-77
- Rosenfeld, J. P., Cantwell, B., Nasman, V. T., Wojdac, V., Ivanov, S., & Mazzeri, L. (1988). A modified, event-related potential-based guilty knowledge test. *International Journal of Neuroscience*, *42*, 157-161.

- Rosenfeld, J. P., Rao, A., Soskins, M., & Miller, A. R. (2002). P300 scalp distribution as an index of deception: control for task demand. *Journal of Credibility Assessment and Witness Psychology*, 3, 1-22.
- Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2008). Time course and robustness of ERP object and face differences. *Journal of Vision*, 8, 1-18.
- Rugg, M. D. & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Science*, 11, 251–257.
- Ryan, J. D., Althoff, R. R., Whitlow, S., & Cohen, N. J. (2000). Amnesia is a deficit in relational memory. *Psychological Science*, 11, 454-461.
- Ryan, J. D., Hannula, D. E., & Cohen, N. J. (2007). The obligatory effects of memory on eye movements. *Memory*, 15, 508-525.
- Schwedes, C., & Wentura, D. (2012). The revealing glance: eye gaze behavior to concealed information. *Memory & Cognition*, 40, 642-651.
- Schwedes, C., & Wentura, D. (2016). Relevance of additional input duration and input information provided by a second fixation on familiarity and recollection when recognizing a face. *Manuscript in preparation*.
- Seymour, T. L., & Kerlin, J. R. (2008). Successful detection of verbal and visual concealed knowledge using an RT-based paradigm. *Applied Cognitive Psychology*, 22, 475-490.

- Seymour, T. L., Seifert, C. M., Shafto, M. G., & Mosmann, A. L. (2000). Using response time measures to assess “guilty knowledge”. *Journal of Applied Psychology, 85*, 30-37.
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison Wesley.
- Unema, P. J. A., Pannasch, S., Joos, M., Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition, 12*, 473-494.
- Verschuere, B., Crombez, G., Degrootte, T., & Rosseel, Y. (2010). Detecting concealed information with reaction times: Validity and comparison with the polygraph. *Applied Cognitive Psychology, 24*, 991-1002.
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*, 441-517.

Author note

Charlotte Schwedes and Dirk Wentura, Saarland University, Faculty of Behavioural Sciences, Department of Psychology, Germany.

Correspondence concerning this article may be addressed to Charlotte Schwedes, Saarland University, Faculty of Behavioural Sciences, Department of Psychology, Building A2.4, D-66123 Saarbrücken (e-mail: c.schwedes@mx.uni-saarland.de).

Acknowledgments

We would like to thank Ullrich Ecker, Regine Bader and Carolyn Wu for their helpful comments on an earlier version of this article as well as Franziska Ebert, Annalena Frey and Annabelle Wallé for data collection.

Appendix

Table A1

List of items the participants interacted with during their mock crime and which are used as probes in the CIT.

	Mock Crime A	Mock Crime B
1	chain of lights	desk light
2	rubbish bin	plastic box
3	garden chair	stool
4	vase	jug
5	pencil case	document wallet
6	permanent marker	white-out
7	diary	index cards
8	hole punch	stapler
9	lever arch folder	folder
10	printer	tape deck
11	file folder box	document tray
12	mini treasure chest	metal tea caddy
13	sticky tape	glue stick
14	Eiffel Tower (keychain)	duck soft toy (keychain)
15	handheld game console	calculator
16	cash box	cash box

Table A2

Frequency and percentage of trials, separate by Display Type, carried forward for analyzes in Experiment 1.

		Number of trials that incurred in the data aggregation						Mean Nr. of trials
Display Type		1	2	3	4	5	6	
Frequency	Target	2	1	2	6	10	13	4.8
	Probe		3	3	4	6	18	5.0
	Irrelevant		2	4	4	11	13	4.9
%	Target	5.9	2.9	5.9	17.6	29.4	38.2	-----
	Probe		8.8	8.8	11.8	17.6	52.9	-----
	Irrelevant		5.9	11.8	11.8	32.4	38.2	-----

Table A3

Means (standard deviations in parentheses) of the duration of the first three fixations as well as the percentage of received first, second and third fixations for Experiment 1, as a function of stimulus type.

	Fixation 1			Fixation 2			Fixation 3			Total fixation duration	
	<i>M</i>	<i>SD</i>	%	<i>M</i>	<i>SD</i>	%	<i>M</i>	<i>SD</i>	%	<i>M</i>	<i>SD</i>
Target-Display											
Targets	267	(71)	100.0	392	(148)	86.9	341	(170)	51.0	1111	(439)
Filler	241	(39)	100.0	340	(115)	77.0	271	(83)	29.4	704	(187)
Probe-Display											
Probes	247	(69)	100.0	414	(146)	81.0	329	(156)	34.5	852	(205)
Filler	269	(64)	100.0	352	(122)	80.4	289	(84)	32.3	770	(183)
Irrelevant-Display											
Irrelevants	239	(63)	100.0	343	(166)	81.8	279	(109)	37.9	828	(230)
Filler	252	(50)	100.0	367	(127)	83.3	272	(58)	40.7	787	(187)

Note. Regarding the mean fixation durations one should keep in mind that the materials used for the fillers were not counterbalanced, and that the aggregate durations of the third fixations are based on fewer trials.

Table A4

Means (standard deviations in parentheses) of the duration of the first three fixations and the total fixation duration as well as the percentage of received first, second and third fixations for Experiment 2a and Experiment 2b, as a function of stimulus type.

		Fixation 1			Fixation 2			Fixation 3			Total fixation duration	
		<i>M</i>	<i>SD</i>	%	<i>M</i>	<i>SD</i>	%	<i>M</i>	<i>SD</i>	%	<i>M</i>	<i>SD</i>
Exp. 2a	Probes	216	(41)	99.8	575	(203)	91.6	647	(199)	68.5	1942	(349)
	Irrelevants	217	(33)	99.8	497	(148)	94.3	594	(153)	76.7	2010	(301)
	Targets	211	(42)	99.8	632	(226)	89.4	633	(210)	66.6	1831	(292)
	Fillers	215	(34)	99.8	574	(203)	91.2	556	(144)	70.1	1874	(241)
Exp. 2b	Probes	209	(41)	99.3	533	(164)	91.8	657	(297)	78.2	1956	(403)
	Irrelevants	215	(43)	99.5	492	(117)	93.5	610	(192)	82.7	1995	(354)
	Targets	209	(42)	99.7	614	(251)	93.2	559	(186)	78.2	1912	(356)
	Fillers	200	(32)	99.9	498	(132)	93.2	606	(185)	81.2	1953	(357)

Note. Regarding the mean fixation durations one should keep in mind that the materials used for targets and fillers were not counterbalanced, and that the aggregated durations of third fixations are based on fewer trials.