The revealing glance: Eye gaze behavior to concealed information

Charlotte Schwedes · Dirk Wentura

Abstract We explored the usefulness of eye fixation durations as a dependent measure in a concealed knowledge test, drawing on Ryan, Hannula, and Cohen (2007), who found eye fixations on a familiar face to be longer than fixations on an unknown face. However, in their study, participants always had to select the known face out of three faces; thus, recognition and response intention could not be differentiated. In the experimental phase of our experiment, participants saw six faces per trial and had to select one of them. We had three conditions: In the first, one of the six faces was a known face, and the participants had to conceal that knowledge and select another face (concealed display); in another, one of the six faces was a known face, and the participants had to select that face (revealed display); or finally, all six faces were unknown, and participants had to select any of the six faces (neutral display). Using fixation durations as the dependent measure, we found a pure and early recognition effect; that is, fixations on the concealed faces (known but not selected) were longer than fixations on the nonselected unknown target faces in the neutral display. In addition, we found a response intention effect; that is, fixation durations on the selected known faces were longer than those on concealed faces (known but not selected).

Keywords Eye movements · Concealed information test · Guilty knowledge test · Fixation duration · Familiarity · Recollection

C. Schwedes (✉) · D. Wentura
Department of Psychology, University of Saarland,
Campus, A2 4,
66123 Saarbrücken, Germany
e-mail: c.schwedes@mx.uni-saarland.de

In recent years, several studies have explored the relationship between recognition of faces and gaze fixations (e.g., Althoff & Cohen, 1999; Hannula et al., 2010; Ryan, Hannula, & Cohen, 2007; Stacey, Walker, & Underwood, 2005). The leading question has been whether prior experience with a face determines subsequent eye movement behavior.

Most importantly for the present research, Ryan et al. (2007) conducted an eye-tracking study in which participants saw three-face displays, with either all of the faces being unknown to participants or one of the faces being a known face. The participants' task was to select the known face in a display. Thus, in the latter condition participants had to select the known face, whereas in the former condition they had to select any one of the three unknown faces. The important result for the present study was the finding of a recognition effect; that is, the known selected faces were fixated for a longer duration than were the unknown selected faces. This effect was even found for the duration of first fixations, indicating fast recognition.

With this finding, Ryan et al. (2007) made an important contribution to the field of indirect memory diagnostics, as their findings suggested that fixation duration (everything else being equal) reveals recognition—that is, a hint as to whether or not someone already knows a certain face. Seen from this perspective, it is interesting to compare the paradigm of Ryan et al. to the best-known paradigm of indirect memory diagnostics—that is, the *concealed information test* (CIT; e.g., Farwell & Donchin, 1991; Langleben et al., 2002; Lykken, 1959, 1960, 1974, 1998; Seymour, Seifert, Shafto, & Mosmann, 2000; Verschuere, Crombez, Degrootte, & Rosseel, 2010; see also the recent volume edited by Verschuere, Ben-Shakhar, & Meijer, 2011), also referred to as the *guilty knowledge test*. In the "classical" CIT, participants are confronted with multiple-choice

questions concerning crime-related details, each containing one correct answer and several foils. For example, "Was the getaway car a (1) red Ford, (2) yellow Toyota, (3) pink Honda, (4) gray Chevy, or (5) white Plymouth?" (Farwell & Donchin, 1991). The participant is instructed to respond "No" to each answer while physiological measures (e.g., heart rate, skin conductivity) are registered (see e.g., Ben-Shakhar & Dolev, 1996; Bradley, MacLaren, & Carle, 1996; for a review, see MacLaren, 2001). Thus, in some sense this test is a variant of what is known as a lie detection test, with the aim of detecting suspects' knowledge about a crime instead of detecting lying.

In more recent versions of the CIT, the underlying logic is somewhat different. Participants are typically confronted with three different kinds of stimuli: crime-related stimuli (probes; i.e., stimuli known by the "guilty" person, but unknown by the "innocent" person), known but not crime-related stimuli (targets), and unknown stimuli (neutral items). The participants' task is to discriminate between targets and the remaining stimuli by, for example, keypresses. Thus, the "innocent" person is confronted with the simple task of discriminating between known (i.e., the targets) and unknown (i.e., probes and neutral items) stimuli, because the probes are indistinguishable from the neutral items. A "guilty" person, however, has a familiarity signal for targets as well as for probes, and must discriminate within the set of known items. On the basis of this fact, the assumption of the modified CIT is that only culprits will show response patterns that consistently deviate for probes as compared to neutral items.

Commonly used dependent measures in this modified CIT variant are event-related potentials (e.g., Allen, Iacono, & Danielson, 1992; Farwell & Donchin, 1991; for a review, see Rosenfeld, 2011) and reaction times (e.g., Seymour & Kerlin, 2008; Seymour et al., 2000; Verschuere et al., 2010). Farwell and Donchin, for example, could differentiate between probes and neutral items for "guilty" participants because the probes elicited an event-related potential in the EEG known as the P300. For "innocents", there were no such differences. Other research groups, such as Seymour et al. (2000), used reaction times as their dependent measure and observed slower reactions to probes than to neutral items for "guilty" participants. This different reaction pattern did not occur for "innocents," because the probes and the neutral items were not differentiable for them.

In the CIT, participants are instructed to conceal their knowledge while these indirect indicators of knowledge are measured. That is, explicit responses are in opposition to possible automatic cognitive and/or bodily reactions. In contrast, Ryan et al. (2007) instructed their participants to pick out the known face if there was one. In our study, we aimed to extend the findings of Ryan et al. by combining

their basic idea of using eye fixations as an indirect indicator of recognition with the CIT paradigm.

The potential advantages of this approach are twofold. First, we could extend the research introduced by Ryan et al. (2007) by disentangling the effects of recognition and of preparing to select an object, an effect labeled the *response intention effect* by Ryan et al. Those authors argued that the fixation duration of the selected known face was driven by recognition processes and the response intention effect, whereas the unknown selected face was driven only by the response intention effect. However, the crucial point was that the differences in the fixation durations of the selected known and selected unknown faces could not definitely be attributed only to the recognition effect. It is possible that the response intention effect is not the same for a selected known face (which one knows immediately should be selected) as for a selected unknown face (which is potentially not chosen until all of the presented faces have been inspected). Thus, to relate the differences in fixation durations exclusively to a recognition effect, a condition would be needed in which a known face was not be selected. The fixation duration on this known face could then be compared to the duration on an unknown face that was also not selected.

The second advantage of a modified paradigm was that with this modification, we created a new version of the concealed information test. We were able to explore the possibility of using eye fixations as an indicator of concealed knowledge (which—to the best of our knowledge—had not been done before; see, e.g., the volume by Verschuere et al., 2011).

Overview

To combine the paradigm of Ryan et al. (2007) with the paradigm of the CIT, we presented three different display types, each containing a photo lineup composed of six faces. For each lineup, we defined one face to be the target and the other five faces to serve as distractors. Whereas the distractors were always unknown, in two-thirds of the displays the target face was known to the participants: In the first third, the known target was a face that had been introduced to the participants as their friend (concealed display). In the second third, the known target was a face that had been introduced as their foe (revealed display). In the final third of the displays, the target was unknown (neutral display).

Participants were instructed to identify their foes and not to betray their friends. More precisely, if presented with a lineup containing the face of a foe (revealed display), they were instructed to choose the foe, but when presented with a lineup containing the face of a friend (concealed display), they had to conceal the knowledge about their friend and

choose one of the unknown faces. If all presented faces were unknown (neutral display), they were to select any of the six faces. Due to this manipulation, we obtained fixation durations for unknown and not selected faces (i.e., the target faces in the neutral displays), for known but not selected faces (i.e., the target faces in the concealed displays), and for known and selected faces (i.e., the target faces in the revealed displays). Thus, in contrast to Ryan et al. (2007), due to the contrast of the concealed and neutral displays, as well as to the contrast of the concealed and revealed displays, we were able to disentangle the recognition and response intention effects. Thus, if the different fixation durations between selected known and selected unknown faces in Ryan et al.'s study were caused by an obligatory recognition effect, we should find longer fixation durations for the target faces in the concealed condition than to the target faces in the neutral condition.

Method

Participants

A total of 37 undergraduate students (29 women, 8 men) from Saarland University took part in the experiment in exchange for course credit.[1] The median age was 21 years (ranging from 19 to 37 years). All had normal or corrected-to-normal vision and were native speakers of German.

Design

Essentially, we had an incomplete 3 (Display Type: concealed, revealed, and neutral displays) × 3 (Face Type: target, selected distractor, and nonselected distractor) within-subjects design. Participants were presented with three different display conditions: (1) The target face was known, but the participant had to conceal it and select one of the five distractor faces (concealed displays), (2) The target face was known and the participant had to select it (revealed display), or (3) the target face as well as the distractor faces were unknown to the participant, and the task was to select any one of the six faces (neutral display). Each display comprised two different face types: one target face and five distractor faces. Additionally, for concealed and neutral displays there was a quasi-experimental variation of nonselected distractors versus selected distractors. Given the instruction to identify foes (i.e., to select the target and not a distractor), this differentiation was missing for revealed displays.

We used a counterbalancing scheme with three sets of lineups and three random samples of participants to ensure that each target face fulfilled the roles of concealed, revealed, and neutral target equally often.

Given our design, the target face in the revealed condition was known and selected; in the concealed condition, it was known but not selected; and in the neutral condition, the target face was both unknown and not selected (due to the exclusion of trials in which the arbitrary neutral target was selected by chance).

The focal hypothesis referred to the comparison between the mean fixation durations of the concealed targets and the neutral targets (i.e., the recognition effect). Given a total sample size of $N = 37$ and α set to .05 (two-tailed), a medium-sized effect ($d = 0.5$) as defined by Cohen (1977) could be detected with a probability of $1 - \beta = .84$ (the power calculations were conducted using G*Power 3; Faul, Erdfelder, Lang, & Buchner, 2007).

Materials

The materials consisted of 108 colored images of German students' faces (54 men and 54 women), taken from a front view or a slight side view, and with a neutral or slightly smiling expression. All of the faces were placed against a uniform gray background. The images measured 174×191 pixels. Overall, we had 18 lineups of six similar faces (i.e., they had the same gender and were similar in their hair color, hair style, and accessories like caps, glasses, or piercings), nine composed of women and nine of men. In each lineup, we randomly selected a target face; the remaining faces of the lineup served as distractors. We created three lists—A, B, and C—of six lineups each (three composed of women and three of men) 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 0.01º. The stimuli were presented with a Windows-based computer on a 17-in. monitor with a viewing distance of 64 cm. The parameters for fixation detection were set to the default values of the eye-tracking software BeGaze (the maximal dispersion value was set to 100 pixels and the minimum fixation duration to 80 ms).

Procedure

The experiment consisted of three phases: a study phase, an experimental phase, and a follow-up test. For each participant, the six targets of one lineup list (A, B, or C) served as faces that were introduced as friends, whereas the six targets of another list served as foes.

---

[1] The data of 10 further participants were defective, due to a loss of eye positions in the experimental phase and to the associated data loss.

In the study phase, participants familiarized themselves with the friends and the foes. The friends' and foes' faces were presented centrally one by one, in a randomized order, while the faces of all six friends were constantly presented in the upper right corner of the display, and the faces of the foes were presented in the upper left corner. Participants had to categorize the centrally presented faces as friends (by pressing the "X" key on a standard keyboard) or foes ("M" key). All 12 faces were presented three times in random order. Error feedback was given in the case of incorrect categorization.

In the subsequent blocks of trials, this procedure was repeated without the faces present in the upper corners of the display. Progression to the experimental phase required completion of at least a further three blocks of trials, with the last block performed without error. If an error happened in the last block, the participants had to pass through one (or more) additional block(s) until no error occurred.

In the experimental phase, the eye movements of participants were recorded. First, the standard nine-point calibration procedure was administered. Then, participants were instructed to imagine that their friends and foes had been involved in a little brawl and were now at a police station. They had to identify each of them in a six-person lineup. The task of the participants was to identify their foes correctly whenever the lineup included one of them (revealed displays), but to protect their friends. Thus, in the case that the lineup included one of their friends (concealed displays), they were instructed to select one of the distractor faces. If no known face was presented in the lineup (neutral displays), participants were instructed to select any face out of the six unknown faces. The participants were also told to behave inconspicuously, so as not to stand out because of different behaviors to the different display types. To familiarize the participants with the procedure in the experimental phase, they had to pass through three practice trials. Each practice trial contained five faces of women and one face of a man. The participants had to identify the man.

Each lineup trial started with a central 500-ms fixation cross, followed by a 50-ms blank display. The participants had to fixate the fixation cross to start the next display (i.e., whenever a drift correction was needed, it could be undertaken at the time of the fixation cross). Then the six faces of a lineup were presented, arranged in a circle. The participants had to select one of the six faces (according to the instructions; see above) and memorize its position. After 7 s, the faces were replaced by masks and the cursor appeared in the middle of the display. Participants had to click with the mouse on the mask that had replaced the selected face. After the response, a blank display was presented again for 50 ms before the next trial started (see Fig. 1). There were six trials each with a concealed, a revealed, and a neutral target face, presented in randomized order. In each lineup, the positions of the faces were selected at random.

Subsequently, to check participants' face knowledge, all concealed and revealed displays were presented again in the follow-up test. The participants' task now was to correctly identify the known face in each lineup. The subsequent display contained the known face of the previous lineup, and the participants had to indicate by a mouse click whether it was a friend's or a foe's face.

At the end of the experiment, participants filled in a questionnaire to check whether they knew any of the faces preexperimentally (which was not the case for any of the participants) and whether they had tried to make use of any strategies in the experimental phase.

Fig. 1 Example of a trial sequence in the experimental phase. (The face stimuli were blurred for this figure.)

500 ms

50 ms

7000 ms

until response

50 ms
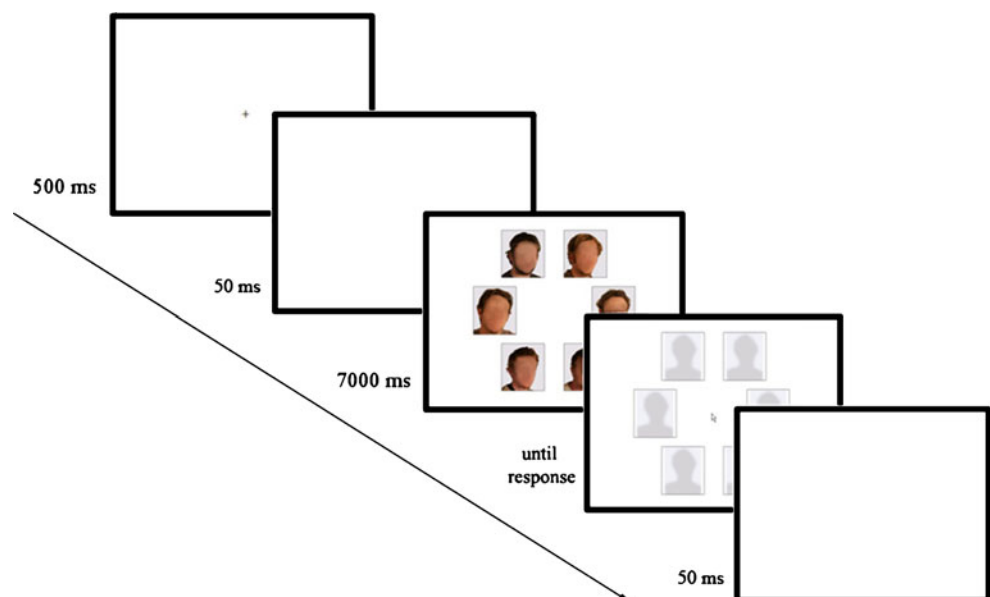
Results

Trials of the experimental phase were excluded from analyses if the participants had responded to the target in a concealed display (2.7% of all concealed-display trials) or had not responded to the target in a revealed display (7.2% of all revealed-display trials). Trials were also excluded if the participant responded incorrectly to the corresponding target in the follow-up test (1.1% of all concealed- and revealed-display trials). In order to match the neutral target condition to the concealed target condition, we also excluded trials in which the arbitrary neutral target had been selected by chance (9.9% of all neutral displays). This last point ensured that the fixation durations to targets in the neutral condition were not biased by response intention. According to these rules, across all participants a total of 49 out of 666 trials (7.4%) were excluded. The data were aggregated using the arithmetic means.[2]

Our results section has three parts: First, we report analyses of the total fixation durations of all stimuli (throughout the 7-s presentation) as the dependent variable. Second, we report analyses for single-fixation durations of the stimuli for the first three fixations. Third, we report preliminary data on the power of our procedure to identify "liars."

Total fixation duration

The means of the total fixation durations as a function of face and display type are depicted in Fig. 2. Because of the incomplete 3 (Display Type) × 3 (Face Type) design, we conducted two separate multivariate analyses of variance (MANOVAs).

The first analysis was a 2 (Display Type: concealed vs. neutral) × 3 (Face Type: target vs. nonselected distractor vs. selected distractor) MANOVA. There, a significant effect of face type, $F(2, 35) = 26.89$, $p < .001$, $\eta_p^2 = .61$, was qualified by a significant interaction between display type and face type, $F(2, 35) = 19.52$, $p < .001$, $\eta_p^2 = .53$.

We used orthogonal Helmert contrasts on the face type variable to test our specific hypotheses about the interaction effect. The first orthogonal 2 (Display Type: concealed vs. neutral) × 2 (Face Type: target vs. distractor) interaction contrast was significant, $F(1, 36) = 35.89$, $p < .001$, $\eta_p^2 = .50$. As can be seen in Fig. 2, the interaction effect is mainly due to the different target fixation durations in concealed and neutral displays, $t(36) = 6.16$, $p < .001$. This difference could be explained by the recognition effect for known faces. The second orthogonal 2 (Display Type: concealed

vs. neutral) × 2 (Face Type: selected distractor vs. nonselected distractor) interaction contrast was not significant, $F(1, 36) = 3.36$, $p = .08$.[3]

To analyze the differences in fixation durations for the targets and nonselected distractors for all three display types, we conducted a second MANOVA—that is, a 3 (Display Type: revealed vs. concealed vs. neutral) × 2 (Face Type: target vs. nonselected distractor) analysis. There was a significant effect of face type (targets vs. nonselected distractors), $F(1, 36) = 67.86$, $p < .001$, $\eta_p^2 = .65$. This main effect was qualified by a significant interaction between face type and display type, $F(2, 35) = 35.46$, $p < .001$, $\eta_p^2 = .67$.

We conducted two planned Helmert contrasts: first, neutral displays versus concealed/revealed displays, and second, concealed versus revealed displays. There was a significantly larger difference between the face types (i.e., target vs. distractor) for the concealed/revealed displays (combined) than for the neutral displays, $F(1, 36) = 69.12$, $p < .001$, $\eta_p^2 = .66$. Second, the difference between the face types in the revealed displays was also significantly larger than the same difference in the concealed displays, $F(1, 36) = 35.83$, $p < .001$, $\eta_p^2 = .50$ (see Fig. 2). This effect between the concealed and revealed displays shows the response intention effect in total fixation durations. In line with this reasoning, the fixation duration for targets was longer in the revealed display as compared to the concealed display, $t(36) = 6.01$, $p < .001$ [$t(36) = -3.57$, $p < .01$, for the nonselected distractors; see above]. Finally, planned comparisons between the targets and nonselected distractors showed significant differences for the revealed displays, $t(36) = 8.66$, $p < .001$, as well as for the concealed displays, $t(36) = 6.07$, $p < .001$, but not for the neutral displays, $t(36) = -1.28$, $p = .21$.

In sum, our analyses of total fixation durations revealed the following results: In the first MANOVA (disregarding the revealed displays), there was a significant Display Type × Face Type interaction that was mainly due to longer fixation durations for the targets in the concealed as

---

[2] To ensure that the following results were not based on outliers, we also conducted data analyses using individual medians for the data aggregation. The results were essentially the same as those subsequently reported.

[3] There were in fact no differences in the fixation durations between the selected distractors of the concealed and neutral displays, $t(36) = -0.21$, $p = .83$. The comparison between the nonselected distractors of the two display types, however, showed a significant effect, $t(36) = -5.49$, $p < .001$, with shorter fixation durations for the nonselected distractors in the concealed displays. This effect could be explained by the limited total fixation duration for a display: If participants were looking longer at the target face in a concealed display, there was less viewing time available for the other faces in that display. It might be argued that due to this dependency, our focal Display Type × Face Type interaction test was biased in favor of our hypothesis. However, if we adjust all mean fixation durations of the nonselected distractors of the concealed displays to the level of the mean fixation durations of the nonselected distractors of the neutral displays, by adding the overall mean difference between these conditions as a constant, the overall Display Type × Face Type interaction remained significant, $F(2, 35) = 15.61$, $p < .001$, $\eta_p^2 = .47$.

Fig. 2 Total fixation durations (in milliseconds) for the three face types, separately for the concealed, neutral, and revealed displays. Error bars are 95% within-subjects confidence intervals for the interaction effect of the first analysis (Jarmasz & Hollands, 2009)
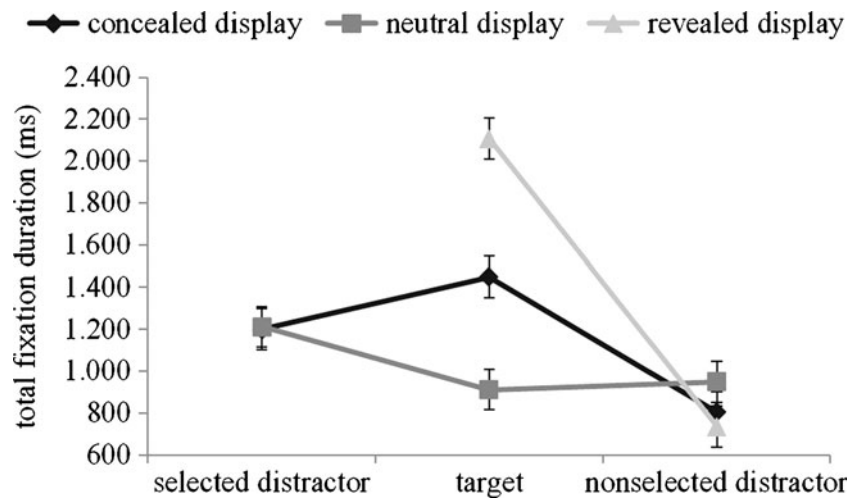
compared to the neutral condition. This difference could be explained as the recognition effect for concealed knowledge. The second MANOVA (disregarding selected distractors) showed a significant Display Type × Face Type interaction as well, which was mainly due to the different fixation durations for the target types. The targets in the neutral condition had shorter fixation durations than did the known (i.e., concealed and revealed) targets, and the targets in the revealed condition had longer fixation durations than did the concealed targets. The latter effect could be explained by the response intention effect. In addition, the difference between the targets and nonselected distractors was significant for the concealed and revealed condition, but not for the neutral condition.

Fixation duration for the first three fixations

To analyze how early the recognition effect appeared, we conducted two further MANOVAs analogous to the ones reported above, but now including the additional factor Fixation (first vs. second vs. third) and employing the average duration of single fixations as the dependent variable.

We constrained the analyses to the first three fixations because each of the six images in a display obtained an average of 3.44 (SD = 1.63) fixations. Thus, including the fourth fixation would have already posed a severe missing data problem. The percentages of images that got a first, second, or third fixation can be seen in Table 1. For the sake of brevity, we report only results for effects including the factor Fixation. Note that each result—except one—that is based on values collapsed across the first three fixations is essentially the same as the corresponding result in the analyses for total fixation times reported above. The exception will be noted at the end of this part of the Results section. The mean fixation durations for the first three fixations, presented separately for face and display types, can be seen in Fig. 3.

The first analysis was a 2 (display type: concealed vs. neutral) × 3 (face type: target vs. nonselected distractor vs. selected distractor) × 3 (fixation: first vs. second vs. third) MANOVA. The triple interaction missed the conventional level of significance, $F(4, 33) = 2.28$, $p = .08$. However, since the focus was on whether the recognition effect appeared already in the first fixation, we additionally

Table 1 Percentages of received fixations (first, second, and third) for each face, reported separately for the display types and face types

		Face Type		
		Selected Distractor	Target	Nonselected Distractor
Fixation 1	Concealed display	99.0	99.5	99.0
	Neutral display	100.0	100.0	100.0
	Revealed display		100.0	98.8
Fixation 2	Concealed display	92.7	97.0	89.7
	Neutral display	99.3	94.0	96.0
	Revealed display		99.5	84.6
Fixation 3	Concealed display	78.2	80.7	59.6
	Neutral display	92.0	79.8	81.2
	Revealed display		92.3	54.9

Fig. 3 Fixation durations (in milliseconds) for the first three fixations (from left to right), separately for the various display and face types. Error bars are 95% within-subjects confidence intervals for the interaction effect of the first analysis (Jarmasz & Hollands, 2009)
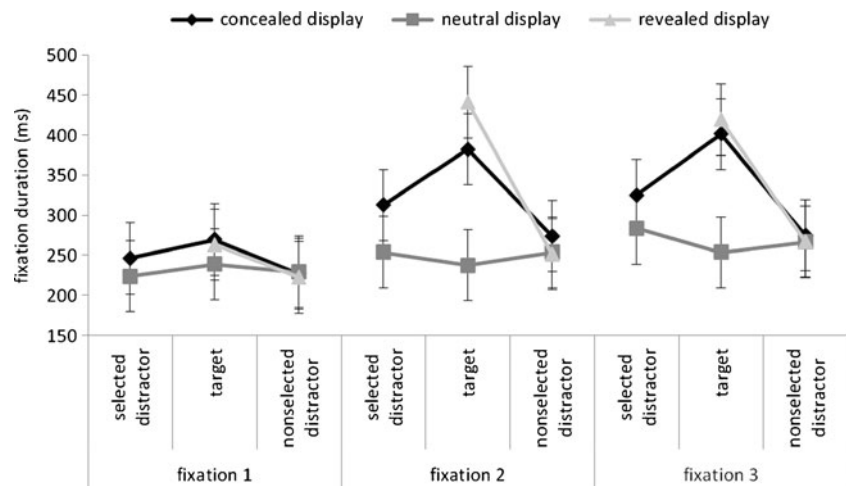
calculated Helmert contrasts for the Fixation factor with first versus second/third fixation as the first contrast and second versus third fixation as the second one. The triple interaction concerning the first Helmert contrast (first vs. second/third fixation) was significant, $F(2, 35) = 5.59$, $p < .01$, $\eta_p^2 = .24$, whereas the interaction concerning the second one (second vs. third fixation) was not, $F(2, 35) < 1$, n.s. Planned comparisons between the targets of the concealed and neutral displays showed no differences for the first fixation, $t(36) = 1.55$, $p = .13$, but did reveal significant differences for the second and third fixations, $t(36) = 5.55$, $p < .001$, and $t(36) = 4.06$ $p < .001$, respectively. Thus, in contrast to the results of Ryan et al. (2007), the recognition effect could clearly be detected in the durations of the second and third fixations but not in the duration of the first fixation.

In the second 3 (display type: revealed vs. concealed vs. neutral) × 2 (face type: target vs. nonselected distractor) × 3 (fixation: first vs. second vs. third) MANOVA, the overall triple interaction was significant, $F(4, 33) = 6.77$, $p < .001$, $\eta_p^2 = .45$. Again, the significance of the overall interaction was especially due to the first Helmert contrast (first vs. second/third fixation), $F(2, 35) = 12.43$, $p < .001$, $\eta_p^2 = .42$ [$F(2, 35) < 1$, n.s., for the contrast of second vs. third fixations]. Planned comparisons between the targets of the revealed and neutral displays showed significant differences for the second and third fixations, $t(36) = 5.92$, $p < .001$, and $t(36) = 3.82$, $p < .001$, but not for the first fixation, $t(36) = 1.66$, $p = .11$. This pattern is similar to the effects between the targets of the concealed and neutral displays reported above. These results show again that the recognition effect could be clearly detected in the durations of the second and third fixations, but not of the first fixation, on a known face (see Fig. 3).

We added one further analysis to settle the question whether our results are really in contrast to those of Ryan et al. (2007), who found the recognition effect already for

the first fixation: Since (a) Ryan et al. had no differentiation comparable to our concealed/revealed distinction and (b) we found no hint of a differentiation of concealed/revealed targets for the first fixation (see Fig. 3), we collapsed over the concealed/revealed conditions and found a significant difference (at least in a one-tailed test) between known (concealed/revealed) targets and unknown (neutral) targets already for the first fixation, $t(36) = 1.91$, $p < .05$ (one-tailed).

Finally, as noted above, there was one exception to the rule that each result of the analyses based on values collapsed across the first three fixations was essentially the same as the corresponding result in the analyses for total fixation time, reported in the first section of these Results. The exception refers to the response intention effect. Whereas the difference between the face types of the revealed displays was significantly larger than the difference for the concealed displays when based on the analysis of total fixation times, the corresponding effect using the collapsed data of the first three fixation missed the conventional level of significance, $F(1, 36) = 2.97$, $p = .09$. The effect remained nonsignificant even when we constrained our analysis to the second and third fixations.

In addition, we analyzed the percentages of stimuli that received a second or third fixation (see Table 1).[4] For second and third fixations, the interactions of face type and display type were significant in both types of analyses (see above), all $Fs(2, 36) > 4.95$, $ps < .05$, all $\eta_p^2 s > .22$. To keep the report concise, we restrict it to differences between targets. A significantly larger number of revealed targets received a second and a third fixation, respectively, as compared to the neutral condition, $t(36) = 2.94$, $p < .01$, and $t(36) =$

---

[4] The percentages of received first fixations for the different display types and face types are all almost 100% (see Table 1). Because of variance limitations, we refrained from analyzing them.

4.25, $p < .001$, respectively. This was not the case for concealed targets, $t(36) = 1.45$, $p = .16$, and $t(36) = 0.25$, n.s., respectively. The corresponding differences between revealed and concealed targets (higher score for revealed targets) missed the conventional level of significance for Fixation 2, $t(36) = 1.62$, $p = .12$, but was significant for Fixation 3, $t(36) = 3.49$, $p < .01$.

In sum, the analysis of the durations of the first three fixations showed no difference in the first fixations between the targets of the revealed and neutral conditions, but did reveal a difference for the second and third fixations. The same pattern was detected for the comparison between the targets of the concealed- and neutral-display conditions, implying no recognition effect before the second fixation. The targets of the revealed and concealed displays did not differ in the first three fixations, so there was no response intention effect in durations within the first three fixations. Only by collapsing the revealed and concealed conditions was there already a difference between targets in the durations of first fixation.

Detecting liars

To see to what extent our eye movement data could be used to differentiate between "guilty" and "innocent" participants, we conducted a further analysis. To this end, we regarded the neutral-display trials as the concealed information test of a virtual "innocent" participant and the concealed-display trials as the concealed information test of a "guilty" (i.e., lying) participant. We checked for each participant, weather the average total fixation duration for the six targets of the given concealed information test was larger than the 95th percentile of the individual distribution of six-item aggregates drawn from the nonselected distractors of the same test.[5] This procedure nominally yielded a false alarm rate of 5% (assuming that the unknown targets were indistinguishable from the nonselected distractors). Factually, 3 out of the 37 "innocent" participants (8.1%) were falsely accused of concealed knowledge. Of the "guilty" participants, the procedure unmasked 24 of the 37 participants (64.9%) as lying.

---

[5] To estimate this individual distribution, we generated a randomized assignment of the four nonselected distractors, separately for each trial, to the placeholders A, B, C, and D. Then we calculated four means for each "innocent" and "guilty" participant by averaging the total fixation durations of the A, B, C, and D distractors of all six neutral trials (i.e., nonselected distractors of the "innocent" participant) and of all six concealed trials (i.e., nonselected distractors of the "guilty" participant), respectively. We repeated this procedure one hundred times in order to estimate distribution parameters based on a total of four hundred mean fixation durations of nonselected distractors.

Discussion

Our study successfully separated a recognition effect from a response intention effect in eye movement behavior during viewing of multiple-face displays with known and unknown faces. Both the recognition of a face and the intention to select a face out of a multiple-face display resulted in increases in the total fixation duration on this face. The effects, however, occurred at different points in time. The recognition effect was already evident in the first fixation (although not robust until the second fixation), whereas the response intention effect did not occur within the first three fixations (with regard to fixation times; but see below) but was observable in the total fixation durations.

The appearance of an early recognition effect is in line with the findings of Ryan et al. (2007). In slight contrast to Ryan et al., we could not find a *robust* recognition effect before the second fixation. This might, however, be a problem of power (we had only 222 trials per condition overall, whereas Ryan et al., 2007, had 324 trials).

It is also possible that the exact point in time at which recognition of a face occurs is affected by the complexity of a scene. In our study, the displays were more complex because they consisted of six rather than three faces (as in the study of Ryan et al., 2007), and in our study all of the faces of a given display were selected for similarity (i.e., they were similar in their hair color, hair style, and whether they wore such accessories as caps, glasses, or piercings). Thus, relative to the materials used by Ryan et al., it was more difficult for our participants to recognize the known faces. It has been shown before that participants are faster to identify distinctive faces than "prototypical" faces (for a review, see, e.g., Johnston & Edmonds, 2009). Eventually, however, by collapsing the data across the revealed and concealed displays, we were able to replicate the finding of a first-fixation difference by Ryan et al. Thus, the recognition effect seems to be a very early effect in eye movement behavior. For this reason, and because of the finding that it seems impossible to conceal the knowledge of a face, we follow Ryan et al. in their notion that the recognition effect on eye movements seems to be obligatory. It remains an open issue whether the longer fixations on known faces are in fact an immediate consequence of the process of memory retrieval (as hypothesized by Ryan et al., 2007) or instead of a delay in the disengagement of attention (see, e.g., Fox, 2004) as a consequence of the recognition of a face.

In contrast to Ryan et al. (2007), in our study participants had to categorize the known face as belonging to either a friend or a foe in order to respond correctly. Therefore, it was possible to separate a recognition effect from a response intention effect for known faces. In fact, we found the response intention effect in the total fixation duration for known faces, with known and "to-be-selected" faces

(revealed target faces) receiving longer total fixation durations than known but "not-to-be-selected" faces (concealed target faces).

It is interesting to note, however, that the differentiation between concealed and revealed target stimuli was not clearly observable within the first three fixations (with regard to fixation time) nor within the first two fixations (with regard to the proportion of targets receiving a first or second fixation). Thus, the most straightforward interpretation of the lack of a difference in eye gaze behavior between concealed and revealed target faces during the first two to three fixations is that participants were not yet able to differentiate between friends and foes.

In this regard, we can relate our results to the differentiation of familiarity-based versus recollection-based memory, known from recognition memory research (see Yonelinas, 2002, for a review). After the identification of the known target face in the display, a recollection process is necessary to differentiate between the two kinds of known targets. The familiarity and recollection processes that are needed to distinguish friends and foes are associated with the activation of two different kinds of components in the classical model of face recognition by Bruce and Young (1986). In this model, the differentiation between known and unknown faces (familiarity effect) is achieved by activation of the corresponding face recognition units. The differentiation between a friend's and a foe's face (i.e., the recollection process) is based on the activation of person identity nodes (PINs), which are units that grant access to semantic face information.

Empirically, familiarity-based recognition is known to have an earlier onset than recollection-based recognition, as revealed by research using event-related potentials (ERPs; see, e.g., Curran, 2000; Rugg et al., 1998; Woodruff, Hayama, & Rugg, 2006). In face recognition, the two ERP effects that are discussed as correlates of familiarity are the N250 and the face N400. The N250 is a negative potential occurring between 200 and 300 ms after stimulus onset; it is the very first ERP that differentiates between known and unknown faces (e.g., Jemel, Schuller, & Goffaux, 2010; Tanaka, Curran, Porterfield, & Collins, 2006). The face N400 occurs around 250 to 500 ms after stimulus onset, and likewise differentiates between known and unknown faces (e.g., Bentin & Deouell, 2000; Webb et al., 2010). Thus, EEG studies suggest that 200–300 ms post-stimulus-onset is the first point in time at which known and unknown faces could be differentiated. The duration of the first fixations in our study—which took on average 266 ms—fits this estimate. However, as long as the PINs are not activated—that is, as long as no recollection process has occurred—there should be no difference between the two kinds of known faces (friends and foes). We can roughly estimate that this process happens during the second or third fixation: The

first reliable difference between friends and foes is the proportion of targets receiving a third fixation, which dropped for neutral and friend targets from second to third fixation but remained at a high level for foe targets. This rough estimate matches the ERP literature, where recollection is typically estimated to occur about 400–800 ms post-stimulus-onset (for reviews, see Johnson, 1995; Rugg, 1995)

The essential point of our work is the finding of a recognition effect in eye gaze behavior for known faces whose recognition was concealed by the participants. In principle, this finding allows for the utilization of eye movement behavior for the detection of concealed information—that is, as an alternative dependent variable in a concealed information test. Although our study was only a first attempt to find a differences in eye gaze behavior between concealed-knowledge items and control items, and was therefore not optimized to detect liars, we found a remarkably high hit rate: detection of about two-thirds of the "lying" participants (while securing a low false alarm rate by design). Relative to the rates in studies using electrodermal measures (for a review, see, e.g., Ben-Shakhar & Elaad, 2003), reaction times (Verschuere & De Houwer, 2011), or ERPs (Rosenfeld, 2011), this hit rate is still comparatively low. It might, however, be enhanced by an advanced experimental procedure, advanced statistics, or simply a higher rate of aggregation.

To strengthen our findings and explanations, further work is needed. It should be clarified which processes contribute to the recognition effect and what factors affect it. The usage of eye movement behavior in the field of indirect memory diagnostics should be further pursued: We need to explore the possibility of faking the recognition effect, the amount of initial exposure to faces, or other items required for the recognition effect to occur, the precision of the differentiation between "innocent" and "guilty" participants, and the influence of nervousness or emotionality on eye movement behavior more generally.

References

Allen, J. J., Iacono, W. G., & Danielson, K. D. (1992). The identification of concealed memories using the event-related potential and implicit behavioral measures: A methodology for prediction in the face of individual differences. *Psychophysiology, 29,* 504–522. doi:10.1111/j.1469-8986.1992.tb02024.x

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.

Ben-Shakhar, G., & Dolev, K. (1996). Psychophysiological detection through the guilty knowledge technique: Effects of mental countermeasures. *Journal of Applied Psychology, 81,* 273–281.

Ben-Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology, 88,* 131–151. doi:10.1037/0021-9010.88.1.131

Bentin, S., & Deouell, L. Y. (2000). Structural encoding and identification in face processing: ERP evidence for separate mechanisms. *Cognitive Neuropsychology, 17,* 35–54. doi:10.1080/026432900380472

Bradley, M. T., MacLaren, V. V., & Carle, S. B. (1996). Deception and nondeception in guilty knowledge and guilty actions polygraph tests. *Journal of Applied Psychology, 81,* 153–160.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77,* 305–327.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Revth ed.). New York: Academic Press.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition, 28,* 923–938. doi:10.3758/BF03209340

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. doi:10.3758/BF03193146

Fox, E. (2004). Maintenance or capture of attention in anxiety-relates biases? In J. Yiend (Ed.), *Cognition, emotion and psychopathology: Theoretical, empirical and clinical directions* (pp. 86–105). New York: Cambridge University Press.

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,* 166. doi:10.3389/fnhum.2010.00166

Jarmasz, J., & Hollands, J. G. (2009). Confidence intervals in repeated-measures designs: The number of observations principle. *Canadian Journal of Experimental Psychology, 63,* 124–138.

Jemel, B., Schuller, A.-M., & Goffaux, V. (2010). Characterizing the spatio-temporal dynamics of the neural events occurring prior to and up to overt recognition of famous faces. *Journal of Cognitive Neuroscience, 22,* 2289–2305.

Johnson, R. J. (1995). Event-related potential insights into the neurobiology of memory systems. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 10, pp. 135–163). Amsterdam: Elsevier.

Johnston, R. A., & Edmonds, A. J. (2009). Familiar and unfamiliar face recognition: A review. *Memory, 17,* 577–596. doi:10.1080/09658210902976969

Langleben, D. D., Schroeder, L., Maldjian, J. A., Gur, R. C., McDonald, S., Ragland, J. D., . . . Childress, A. R. (2002). Brain activity during simulated deception: An event-related functional magnetic resonance study. *NeuroImage, 15,* 727–732

Lykken, D. T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology, 43,* 385–388.

Lykken, D. T. (1960). The validity of the guilty knowledge technique: The effects of faking. *Journal of Applied Psychology, 44,* 258–262.

Lykken, D. T. (1974). Psychology and the lie detector industry. *American Psychologist, 29,* 725–739. doi:10.1037/h0037441

Lykken, D. T. (1998). *A tremor in the blood: Uses and abuses of the lie detector.* New York: Plenum.

MacLaren, V. V. (2001). A quantitative review of the guilty knowledge test. *Journal of Applied Psychology, 86,* 674–683. doi:10.1037/0021-9010.86.4.674

Rosenfeld, J. P. (2011). P300 in detecting concealed information. In B. Verschuere, G. Ben Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 63–89). Cambridge: Cambridge University Press.

Rugg, M. D. (1995). ERP studies of memory. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind* (pp. 132–170). New York: Oxford University Press.

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature, 392,* 595–598.

Ryan, J. D., Hannula, D. E., & Cohen, N. J. (2007). The obligatory effects of memory on eye movements. *Memory, 15,* 508–525.

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. doi:10.1002/acp.1375

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. doi:10.1037/0021-9010.85.1.30

Stacey, P. C., Walker, S., & Underwood, J. D. (2005). Face processing and familiarity: Evidence from eye-movement data. *British Journal of Psychology, 96,* 407–422.

Tanaka, J. W., Curran, T., Porterfield, A., & Collins, D. (2006). The activation of pre-existing and acquired face representations: The N250 ERP as an index of face familiarity. *Journal of Cognitive Neuroscience, 18,* 1488–1497.

Verschuere, B., Ben-Shakhar, G., & Meijer, E. (2011). *Memory detection: Theory and application of the Concealed Information Test.* Cambridge: Cambridge University Press.

Verschuere, B., Crombez, G., Degroote, T., & Rosseel, Y. (2010). Detecting concealed information with reaction times: Validity and comparison with the polygraph. *Applied Cognitive Psychology, 24,* 991–1002. doi:10.1002/acp.1601

Verschuere, B., & De Houwer, J. (2011). Detecting concealed information in less than a second: Response latency-based measures. In B. Verschuere, G. Ben Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 46–62). Cambridge: Cambridge University Press.

Webb, S. J., Jones, E. J. H., Merkle, K., Murias, M., Greenson, J., Richards, T., . . . Dawson, G. (2010). Response to familiar faces, newly familiar faces, and novel faces as assessed by ERPs is intact in adults with autism spectrum disorders. *International Journal of Psychophysiology, 77,* 106–117

Woodruff, C. C., Hayama, H. R., & Rugg, M. D. (2006). Electrophysiological dissociation of the neural correlates of recollection and familiarity. *Brain Research, 1100,* 125–135.

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. doi:10.1006/jmla.2002.2864