Systematic influence of gaze position on pupil size measurement: analysis and correction

Benjamin Gagl · Stefan Hawelka · Florian Hutzler

Abstract Cognitive effort is reflected in pupil dilation, but the assessment of pupil size is potentially susceptible to changes in gaze position. This study exemplarily used sentence reading as a stand-in for paradigms that assess pupil size in tasks during which changes in gaze position are unavoidable. The influence of gaze position on pupil size was first investigated by an artificial eye model with a fixed pupil size. Despite its fixed pupil size, the systematic measurements of the artificial eye model revealed substantial gaze-position-dependent changes in the measured pupil size. We evaluated two functions and showed that they can accurately capture and correct the gaze-dependent measurement error of pupil size recorded during a sentence-reading and an effortless z-string-scanning task. Implications for previous studies are discussed, and recommendations for future studies are provided.

Keywords Pupillometry · Eye movements · Cognitive effort · Video-based eyetracker · Reading · Mindless reading

Introduction

The pupil does not respond only to variations of illumination, but changes in the dilation also reflect cognitive load, arousal, or emotional valence (Beatty, 1982; Partala & Surakka, 2003; Stanners, Coulter, Sweet, & Murphy, 1979). Task-evoked changes of pupil size in response to cognitive demands were documented for experimental manipulations of mental calculation (Hess & Polt, 1964), working memory load (Granholm, Asarnow, Sarkin, & Dykes, 1996; Kahneman & Beatty, 1966; Peavler, 1974; Piquado, Isaacowitz, & Wingfield, 2010), and visual search (Porter, Troscianko, & Gilchrist, 2007). Typically, increased cognitive demands are reflected by an increased dilation of the pupil. The pupil, in general, starts to dilate about 200–300 ms after stimulus presentation and reaches a task-dependent ceiling during stimulus processing (for a review, see Beatty, 1982).

In the domain of visual word recognition, recent studies have revealed that the presentation of isolated words results in a similar pupillary response pattern. Again, the pupil starts to dilate around 200–300 ms after the presentation of a word and reaches a ceiling around 1,200 ms (Briesemeister et al., 2009; Kuchinke, Võ, Hofmann, & Jacobs, 2007; Võ et al., 2008). In contrast and as detailed below, for the processing of whole sentences (with at least several words being presented simultaneously on the screen), a divergent pupillary response pattern was reported (sentences, Just & Carpenter, 1993, and Schluroff et al., 1986; fragments of sentences, Raisig, Welke, Hagendorf, & van der Meer, 2007, 2010). The obvious difference between studies on sentence processing and visual word recognition is the necessity for changes in gaze position during the processing of a series of words, which are presented simultaneously on the screen. In the present study, we will show that changes in gaze position systematically affect the measurement of pupil size, suggesting that the divergent pupillary response during sentence reading is, in fact, a measurement error. Subsequently, an approach that allows correcting for this measurement error will be presented.

In studies on visual word recognition, it has been shown that the higher cognitive effort that accompanies processing of low-frequency words (i.e., words that rarely occur in

B. Gagl (✉) · S. Hawelka · F. Hutzler
Department of Psychology and Center
for Neurocognitive Research, University of Salzburg,
Hellbrunnerstrasse 34,
5020 Salzburg, Austria
e-mail: Benjamin.Gagl@sbg.ac.at

written language) results in an increased dilation of the pupil, as compared with the less effortful processing of high-frequency words (Kuchinke et al., 2007). A similar finding was reported by Briesemeister et al. (2009), who showed an increased dilation of the pupil in response to the higher processing demands of pseudohomophones (i.e., nonwords that are pronounced like an existing word), as compared with the processing of real words. Furthermore, pupil dilation was shown to reflect processes of memory retrieval for words (Võ et al., 2008): In a recognition memory task, an increased dilation of the pupil was found in response to hits, as compared with correct rejections. These studies on visual word recognition reported the typical pupil response with a maximal dilation of the pupil at around 1,200 ms after the presentation of the stimulus. It is important to mention that the above-cited studies relied on the presentation of isolated words, which were centered on the screen and, thus, did not require any eye movements (i.e., changes in gaze position) to perceive the stimuli.

The presentation of whole sentences, in contrast, has been used to investigate, for example, pupil response as an indicator of cognitive demands during syntactic processing. In this context, it is important to state that the reading of whole sentences requires horizontal eye movements (in contrast to the presentation of isolated words in studies of visual word recognition). Schluroff et al. (1986) reported that the processing of syntactically complex sentences resulted in increased dilation of the pupil, as compared with the processing of sentences with less syntactic complexity (see also Just & Carpenter, 1993). In a study by Raisig et al. (2007), participants had to examine fragments of sentences in order to explore behavioral scripts (sequences of daily life activities; e.g., *The boy comes home; has dinner; goes to bed*). Noteworthy is that the authors of the Schluroff et al. and Raisig et al. (2007) studies were quite detailed about the results and plotted the complete time course of the pupil response. Surprisingly, the change of pupil size over time deviated remarkably from the change reported in studies with single-word presentation: Prior to a dilation of the pupil, an initial constriction of the pupil was reported. This unusual constriction was interpreted as possibly reflecting "a habituation to the situation" (Raisig et al., 2007, p. 869). On the basis of this accurately reported data, and taking into account that the participants had to move their eyes horizontally across the screen during the reading of the sentences, one might wonder whether this unusual initial constriction might be induced by changes in gaze position. Manufacturers of video-based eye-tracking systems acknowledge that the measurement of pupil size by their systems is affected by gaze position ("up to 10%"; S-R Research Eyelink-CL Manual, p. 101; see also Scheepers, & Crocker, 2004). Changes in gaze position might exert

their systematic influence on pupil size due to a change in the geometry of the recorded (image of the) pupil: When a participant moves his/her gaze to different locations on the screen, the shape of the recorded pupil alters from a circle to an ellipse (Atchison & Smith, 2000, p. 25). This change in the recorded geometry of the pupil is accompanied by a change in the measured pupil size. An additional source for a measurement error might be the anatomy of the iris (more specifically, the thickness of the iris) that defines the pupil (Jay, 1962). Scheepers and Crocker acknowledged the importance of accounting for gaze position changes during tasks that involve eye movements. They controlled for the influence of gaze position by subtracting the predictions from interindividual multiple-regression analyses with the x- and the y-coordinates of the gaze position as predictors from the raw data. However, they did not provide a systematic investigation and an empirical verification of the issue.

In the present study, we document the influence of gaze position on pupil size. This was realized by measuring the pupil size of an artificial eye with a constant pupil size, the "gaze" of which was directed to different screen positions. Since the measured pupil size of the artificial eye can be influenced only by changes in gaze position, this approach is suitable for approximating a pure, gaze-dependent measurement error. Additionally, we measured participants' pupil size during an effortful sentence-reading task and a control task of minimal cognitive demands—that is, z-string scanning (Nuthmann, Engbert, & Kliegl, 2007; Rayner & Fischer, 1996; Vitu, Oregan, Inhoff, & Topolski 1995). All measurements were assessed with a video-based, corneal reflection eye-tracking system, an eye-tracking technique that is widely used in pupillometry and eye-tracking research. Subsequently, a correction approach will be introduced, which compensates for the influence of gaze position on pupil size measures, thus allowing an unambiguous interpretation of the pupil response.

Method

Participants

Forty-nine (13 male; 18–47 years old; $M = 24$ years) native German-speaking students at the University of Salzburg participated in the present study. All participants had normal vision. Participants with contact lenses or glasses were not included, to prevent unpredictable optical artifacts.

Apparatus

Pupil size and eye movements were recorded from the right eye with an SR-Research Eyelink 1000 tower mount

system with a sampling rate of 1000 Hz. The head of the participants was stabilized with a chin- and a forehead-rest at a distance of 60 cm in front of a 21-in. CRT screen (1,024 × 768 pixel resolution; 120-Hz refresh rate), which was connected to an IBM-compatible PC. The Eyelink 1000 corneal reflection system assesses gaze position changes by measuring the reflection of an infrared illuminator on the cornea, as well as the size of the pupil, with a video camera that is sensitive to light in the infrared spectrum. This method of pupil size and eye movement assessment is widely used in various eye trackers and will allow a generalization of the findings of the present study.

The pupil is assessed in the centroïd mode of the eye tracker, which uses a center-of-mass algorithm. This algorithm detects the pupil area by identifying the number of black pixels and its center on the video image. Importantly, in contrast to methods using ellipse fitting for the measurement of the pupil, this method is hardly affected by noise (S-R Research Eyelink-CL Manual, p. 71).

In the present study, pupil size was measured on the basis of the assessed pupil area. This pupil size measure is similarly affected by gaze position changes in both the horizontal and the vertical directions. Alternatively, for the present investigation of horizontal eye movements during sentence reading, we also could have used a vertical diameter measure. This diameter measure, however, would have prevented a generalization to paradigms that involve gaze position changes in both the horizontal and the vertical dimensions. If the gaze positions changes horizontally, the vertical diameter should not be affected. But if a gaze position change involves a vertical change or even a change in both dimensions, the vertical diameter will be strongly affected (see Atchison, & Smith, 2000, p. 25). Importantly, when the gaze is changing in both dimensions, a correction or interpretation of the results on the basis of a single-dimension diameter measure would be nearly impossible.

The illuminance level in the windowless laboratory was held constant at about 100 lux for all participants. The stimuli were presented in black bold Courier New (14 pt.; with a single letter extending 0.3° of horizontal visual angle) on a white background by the Experiment Builder software (SR Research Ltd., Canada). The luminance of the screen was approximately 8 cd/m².

Materials and procedure

For the sentence-reading task, the Potsdam Sentence Corpus (PSC; Kliegl, Grabner, Rolfs, & Engbert, 2004) with a high variability of syntactic structures was used. The PSC consists of 144 German sentences with a mean length of 7.9 words (overall, 1,183 words). Participants were instructed to read silently for comprehension and had to answer simple comprehension questions after 25% of the sentences. The z-strings for the mindless reading task were constructed by replacing all letters of the PSC words with the letter z—preserving capitalization, punctuation marks, and word boundaries (see Nuthmann et al., 2007; Vitu et al., 1995). Participants were instructed to scan the meaningless z-strings, mimicking reading.

To familiarize the participants with the tasks, 10 practice trials preceded the experimental trials. In both tasks, a short break was provided at halftime. The eye tracker was calibrated by a nine-point calibration routine at the beginning of each task and after every break (calibration error was kept below 0.5° of visual angle). Constant accuracy of eye-tracking measurement was ensured by a fixation check at the beginning of each trial: A fixation point was presented in a distance of 100 pixels from the left side of the monitor and remained on the screen until fixated by the participant (or up to 5 s in a case of insufficient calibration, resulting in recalibration). The minimum duration of the fixation check was 70 ms. After a successful fixation check, the experimental stimuli (sentence or z-strings) were presented. The horizontal position of the sentences or sequences of z-strings was adjusted so that the middle of the first-word/z-string was at the position of the preceding fixation point. Participants terminated stimulus presentation by looking at an "x" presented in the lower right corner of the screen.

Artificial eye model

To realize a condition with constant physical pupil size, an artificial eye model with a pupil diameter of 4.5 mm was constructed (see Fig. 2b). This model of an eye was fixed by a mechanical apparatus, and its orientation was controlled by a goniometer. For a simulation of different horizontal gaze positions, the orientation of the artificial eye model was varied—covering the horizontal extent of the presented sentences (from about -17° to +9° relative to the center of the screen). Since the artificial eye data revealed a drastic change in measured pupil size for the leftmost part of the screen (see below), the following horizontal step sizes (i.e.; saccade length of the artificial eye) were realized: From 70 pixels (about -17°, the leftmost position for which an initial fixation was observed during sentence reading and z-string scanning) to 130 pixels (-13.5°), a step size of 20 pixels was chosen; from 130 pixels onward, a step size of 100 pixels was chosen. This resulted in 11 bins for the measurement of the artificial eye model.

Results

Participants had no problem comprehending the sentences (98% of the comprehension questions were answered correctly). During sentence reading and z-string scanning,

the participants exhibited a comparable mean number of fixations per line [M = 9.1 and 9.9 fixations, respectively; t (48) = 1.1] and a comparable saccade length [M = 7.9 and 7.7 letters, respectively; t(48) = 0.9]. In contrast, the mean fixation duration was significantly prolonged during z-string scanning, as compared with sentence reading [M = 243 and 192 ms, respectively; t(48) = 8.9, p < .001]. We note that this pattern of results closely replicates previous findings from studies in which z-string scanning was compared with sentence reading (e.g., Nuthmann et al., 2007; Vitu et al., 1995).

Pupil dilation

For the analysis of pupil dilation, the continuous measurement of pupil size was segmented into epochs from 70 ms prior to stimulus onset until the end of the trial. Eye blinks within a trial were identified visually, and only the respective blink intervals were removed from analysis. Baseline pupil size was calculated on the basis of a 70-ms prestimulus interval and did not differ between the sentence-reading and the z-string-scanning tasks[M = 1,368 pixels, SD = 345, and M = 1,326 pixels, SD = 293, respectively; t(48) = 0.64]. The change in pupil size (relative to the prestimulus baseline) dependent on horizontal gaze position is plotted in Fig. 1a, separately, for sentence reading, z-string scanning, and the artificial eye. For the z-string and sentence-reading tasks, the bin size of the change in measured pupil size relative to the horizontal gaze position (hereafter, pupil response) was 10 pixels of horizontal gaze position.

For all three data sets, there was a steep increase in measured pupil size from the leftmost position (-17°) to -15°, where a peak could be observed. From this position onward (-15°), the measured pupil size decreased monotonically until it asymptotically stabilized at the rightmost horizontal measurement positions. This observed peak around -15° is presumably a result from the position of the camera relative to the eye. In the present tower mount setup, the camera was positioned slightly to the left of the eye, which constituted an angle of about 18° (5 mm). The position of the camera is responsible for the particular shape of the pupil response. To illustrate, when the gaze is perpendicular to the screen (at about 3° relative to the center of the screen), the shape of the measured pupil is elliptic, due to the left-shifted position of the camera. When now the gaze position changes to the left (i.e., toward the camera), the elliptic shape of the pupil becomes more and more circular—that is, increasingly larger. At about -15°, the measured pupil size is largest, as a result of the (now) perpendicular view of the camera on the eye. Thus, the position of the camera relative to the eye accounts for the particular shape of the pupil response. However, note that the influence of the camera position will be accounted for by the gaze position correction procedures (see below).

In a first analysis, the pupil response during the z-string-scanning task and that obtained from the artificial eye model were compared: The pupil response measured during z-string scanning did not differ from the pupil response obtained by the artificial eye model (at 11 different horizontal gaze positions; one-sample t-tests, all ts < 1). In a second analysis, the comparison of the pupil responses from z-string scanning and sentence reading revealed, from a horizontal gaze position of about -12° onward, an overall larger measured pupil size during sentence reading (all ts > 2.3, all ps < .05) than during z-string scanning .

Interim conclusion

Two findings are of interest. First, the higher cognitive demands of sentence reading, as compared with z-string scanning, were reflected by a larger pupil size during sentence reading. Importantly, an initial steep increase with a subsequent monotonic decrease in measured pupil size was observed not only in both tasks, but also for the artificial eye model. Whereas one might be tempted by a cognitive interpretation of this measured pupil response during sentence reading and z-string scanning, the data from the artificial eye model indicate a systematic measurement error: Although the artificial eye has a constant pupil size, the pupil size as measured by the eye-tracking system changed with its orientation (i.e., gaze position). Thus, the pupil size as measured by a video-based eye-tracking system (independent from the actual size of the pupil) is affected by the horizontal gaze position. Therefore, the measurements of pupil size have to be corrected for horizontal gaze position in order to prevent misinterpretations of the pupil response.

Correction

In the previous section, we showed that horizontal gaze position does have a systematic influence on pupil size as measured by a video-based eye-tracking system (hereafter, measurement error). The "pure" measurement error was revealed by the artificial eye model, which has a constant pupil size, but for which the eye tracker nevertheless reported varying pupil size data dependent on the horizontal gaze position. In the following, we will describe (and predict) this measurement error with a mathematical function that is dependent on the horizontal gaze position. Such a correction function would allow one to predict the magnitude of the measurement error for a specific horizontal gaze position and to subtract this measurement error from experimentally acquired pupil data.

Two such correction approaches will be presented and evaluated below: (1) a correction approach that attempts to utilize the physical parameters of the experimental setting and (2) an alternative approach based on a synthetic function, derived from polynomial curve fitting. These functions will

Fig. 1 **a** Means of uncorrected sentence-reading, z-string-scanning, and artificial eye pupil data, relative to the horizontal gaze position. **b** Means from the artificial eye data and predictions from the physical and synthetic functions. **c** Means of the corrected pupil data during sentence reading (corrected by artificial-eye- and z-string-based synthetic functions) and z-string scanning (corrected by z-string-based synthetic functions). Pupil size in percentage of change was calculated relative to the baseline pupil size
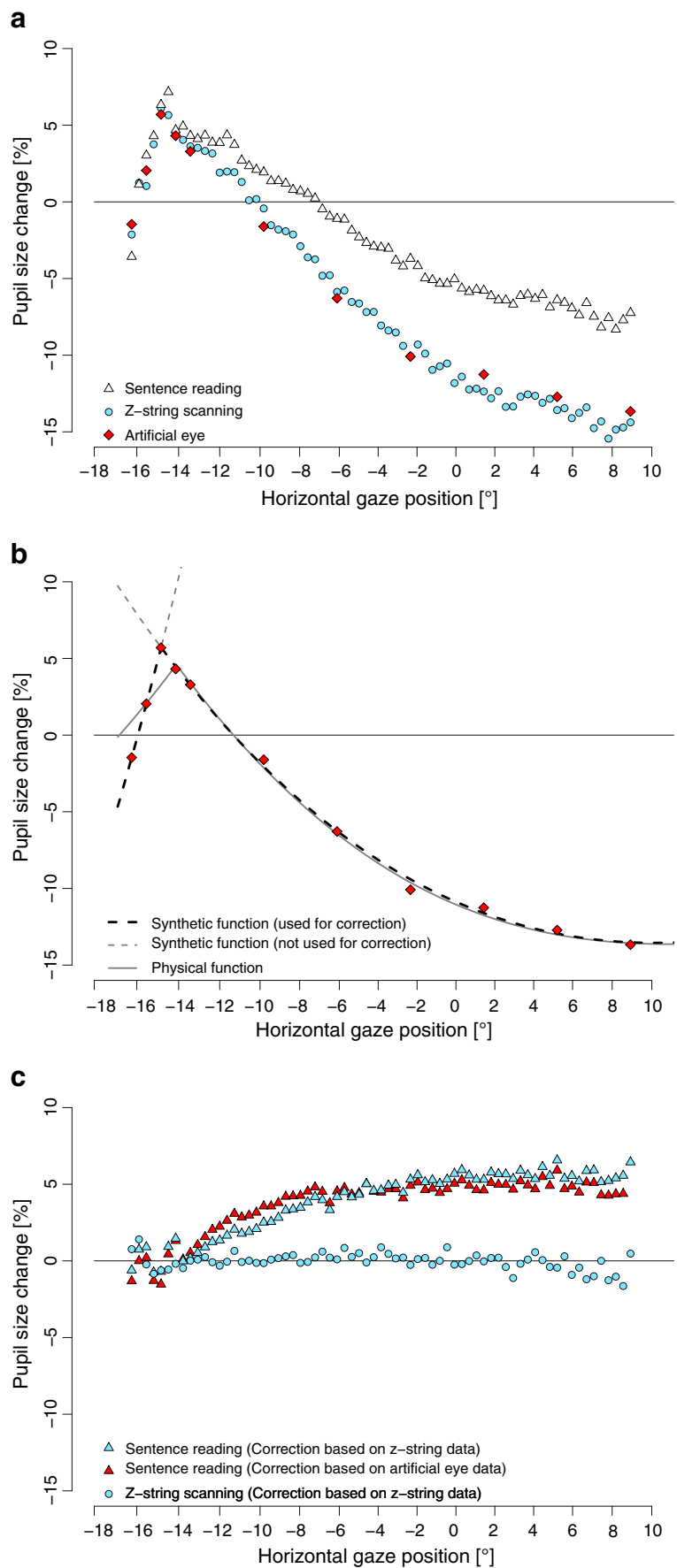
be evaluated with respect to the accuracy by which they can estimate the "pure" gaze-dependent measurement error as approximated by the artificial eye data.

Correction approach A: Physical function

The highly controlled setting of the artificial eye model allows us to describe the physical properties that are probably responsible for the gaze-dependent measure error (such as the visual angle of the gaze relative to the screen center) in a mathematical function (see Fig. 2a). For this formalization, measureable physical properties of the artificial eye and the measurement setting are treated as fixed parameters. The three fixed parameters are (1) the distance from the eye to the perpendicular position on the screen (D_S = 600 mm), (2) the constant pupil size of the artificial eye (radius: P_0 = 2.25 mm), and (3) the horizontal position of the eye relative to the screen (equivalent to the gaze position when the gaze is perpendicular to the screen; G_0 = 290 mm from the left of the screen). Four free parameters are used to fine-tune the physical parameters named above and to substitute physical parameters that are not measureable. In short, the first parameter (F_{P1}) accounts for the exact horizontal position of the artificial eye. The second parameter (F_{P2}) allows adjusting the eyetracker measurement to the properties of the artificial eye's pupil. The third parameter (F_{P3}) accounts for the thickness of the iris. The final parameter (F_{P4}) allows an adaptation of the physical function to the baseline pupil size. A detailed description of the four parameters is provided in the following.

In Eq. 1, the actual horizontal gaze position (measured relative to the left of the screen [G_P]) is transformed to "distance relative to G_0"(i.e., distance to the center of the screen: G_{PC}; positive deviation in both directions). Additionally, the first free parameter (F_{P1}) is introduced, which allows an optimization of the exact horizontal position of the artificial eye:

$$G_{PC} = |G_P - G_0 - F_{P1}|. \tag{1}$$

In the next step, the relative gaze position (G_{PC}, as estimated in Eq. 1) is transformed into a visual angle α relative to G_0 by a trigonometric function taking into account the distance from the eye to the screen (D_S):

$$\alpha = \sin^{-1} \frac{G_{PC}}{\sqrt{G_{PC}{}^2 + D_S{}^2}}. \tag{2}$$

With varying horizontal gaze positions, the assessed shape of the pupil changes from a circle (when the gaze is perpendicular to the camera—i.e., α = -15°) to an ellipse (when the visual angle diverges from -15°; see Fig. 2c). The horizontal axis of this resulting ellipse (P_G) is estimated by a trigonometric function that takes into account the visual

angle (α) and the fixed pupil size of the artificial eye (P_0). A further free parameter (F_{P2}) is introduced that does allow correcting for a potential constant divergence between the fixed pupil size of the artificial eye and the pupil size as measured by the eye tracker, which can be caused by differential anatomical properties of the artificial, as compared with the human, eye:

$$P_G = P_0 * F_{P2} * \cos \alpha. \tag{3}$$

In the next step, the thickness of the iris is implemented, since it additionally influences the assessment of the pupil when the eye is not perpendicular to the camera (see Fig. 2d). The iris is the border of the pupil and does have a certain thickness. When the gaze is not perpendicular to the camera, the walls of the iris (i.e., its thickness) additionally reduce the size of the measured pupil (see Fig. 2d, right panel). To illustrate this, the influence of the thickness of the iris can be compared with the field of vision that one has when looking through a pipe: When the pipe is perpendicular, the openings of the pipe align, and the field of view is maximal. When the pipe is slightly angled, the walls of the pipe move into view. For the angled pipe, the reduction of the overlap of the openings is comparable to the influence of the thickness of the iris on the measurement of the size of the pupil. In the following, we accounted for thickness of the iris by assuming two ellipses of equal size (one in front and one behind the iris; Fig. 2d, dark gray lines; horizontal axis: P_G), which are comparable to the openings of the hypothetical pipe. When the eye is perpendicular to the camera, the two openings (Fig. 2d, left panel) are perfectly overlapping. But when the gaze position changes, the eye does have a specific angle with respect to the camera (denoted by α), and thus, the ellipses are not perfectly aligned anymore (Fig. 2d, right panel). In sum, the overlap (hereafter, intersection area) of the two ellipses allows us to account for the influence of the thickness of the iris on measured pupil size.

First of all, the estimation of the intersection area of the two ellipses was realized by calculating the horizontal distance of the centers of the ellipses (A; see Fig. 2d, right panel). The estimation is based on a trigonometric function determined by the visual angle (α) and the unknown distance of the two ellipses (i.e., the thickness of the iris), which is the third free parameter (F_{P3}):

$$A = F_{P3} * \sin \alpha. \tag{4}$$

In the next two steps, the horizontal and the vertical axes of the intersection area ($E_{Horizontal}$ and $E_{Vertical}$, respectively) are estimated. In Eq. 5, the horizontal axis is estimated on the basis of the distance of the ellipses'centers (A), just as the horizontal axis of the pupil ellipse (P_G):

$$E_{Horizontal} = 2 * P_G - A. \tag{5}$$

Fig. 2 Schematic illustration of the **a** experimental setting, **b** vertical midline cut of the artificial eye model, **c** change in pupil geometry, and **d** influence of the thickness of the iris on the measured pupil size when the gaze is not perpendicular to the camera
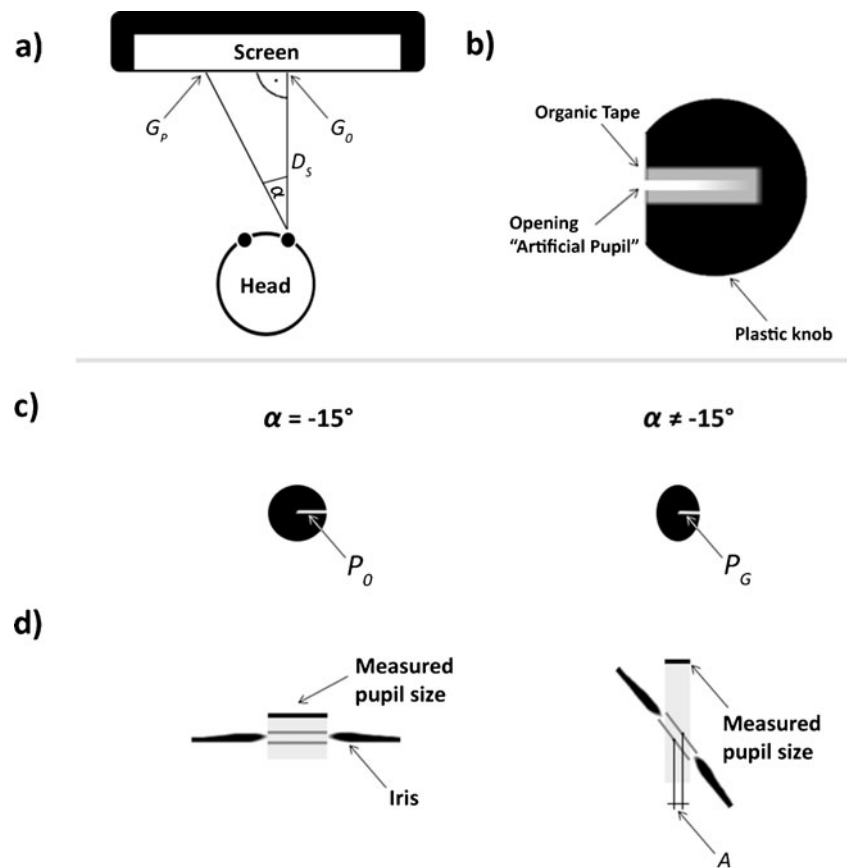
The vertical axis of the intersection area ($E_{vertical}$) is, basically, the distance between the two points at which the ellipses intersect. The distance of the intersection points can be estimated via a binomial formula:

$$E_{Vertical} = \sqrt{4*(P_0*F_{P2})^2 - \frac{A^2*(P_0*F_{P2})^2}{P_G^2}}. \tag{6}$$

In the final step, we estimated the intersection area (E_{GPphys}) as an ellipse—which is asufficient approximation—on the basis of π, the horizontal axis, the vertical axis, and an additive free parameter ($F_{P4}{}^1$) that adapts to the baseline pupil size:

$$E_{GPphys} = \pi*E_{Horizontal}*E_{vertical} + F_{P4}. \tag{7}$$

The physical function, as derived above, can now be used to estimate the measurement error for a specific horizontal gaze position (G_P). We adjusted the four free parameters (F_{P1}–F_{P4}) by a nonlinear least-square fitting approach to achieve the best possible fit of the artificial eye data, which serves as an estimate for the "pure" gaze-dependent measurement error.[2]

The physical function (Fig. 1b, solid gray line) can explain a high proportion of the variance of the artificial eye data (R^2 = .97; mean squared error [hereafter, *MSE*] = 167 pixels). In summary, the physical parameters, which are the basis for the physical function, can relatively accurately account for the gaze-dependent measurement error.

Correction approach B: Synthetic function

The physical function was based on the physical parameters of the measurement setting. An alternative approach would be to use an arbitrary synthetic function, which can be easily designed with the mere aim of a best possible fit to the artificial eye data. A visual inspection of the artificial eye data reveals a relatively steep positive slope to the left of the reversal point (-17° to-15°) but a moderate negative

---

[1] One might wonder whether different luminance levels of the screen would affect the shape of the pupil response. From the similarities of the response from the artificial eye and the human eye, we suggest that the shape of the response should not be affected by different luminance levels. In the physical function, the free parameter F_{P4} would account for such a difference by adjusting the baseline pupil size in accordance with the luminance level of the actual experimental setup. This adaptation of F_{P4} results in a vertical adjustment of the predictions from the physical function to the baseline of the to-be-fitted pupil response data. Thus, one can adapt the physical function for future studies with different experimental displays.

---

[2] The R code used in the present study can be obtained from the corresponding author upon request.

slope to the right of the reversal point (from -15° onward). These different characteristics of the slope can be best accounted for by two independent functions. Eq. 8 specifies the polynomial fitting approach ($E_{GPsynth}$) that was used to estimate the gaze-dependent measurement error:

$$E_{GPsynth} = I + \beta_1 {}^* G_P + \beta_2 {}^* G_P{}^2 + \ldots + \beta_4 {}^* G_P{}^4 \qquad (8)$$

To appropriately capture the nonlinear characteristics of the artificial eye data, we implemented—besides the intercept (I) and the linear term ($\beta_1 {}^* G_P$; first order)—a quadratic term ($\beta_2 {}^* G_P^2$; second order), a higher order cubic ($\beta_3 {}^* G_P^3$; third order), and a quartic term ($\beta_2 {}^* G_P^2$; fourth order). Analogous to the physical function, we submit the actual gaze position (G_P) to the function. I and $\beta_1 - \beta_4$ serve as free parameters, which are adjusted to achieve the best possible fit to the artificial eye data (nonlinear least-square fitting approach). It is important to note that two different sets of parameters were used to fit the artificial eye data to the left and to the right of the reversal point (i.e., two different functions being used to fit the artificial eye data to the left and to the right of the reversal point).

The synthetic function obtained an almost perfect fit of the artificial eye data (Fig. 1b, black dashed line vs. diamonds; $R^2 > .99$ for both functions; $MSE = 15$ pixels; complete formulas are listed in the Appendix). In summary, the synthetic function captures the "pure" gaze-dependent measurement error with higher accuracy than does the physical function. Consequently, for the subsequent correction approach, the synthetic function will be used.

A pragmatic alternative: Correction based on z-string data

The "pure" measurement error as obtained with the artificial eye data was fitted best by a synthetic function, which will be used to estimate the gaze-dependent measurement error to correct the sentence-reading data. The technically challenging assessment of the artificial eye motivated us to provide an alternative, more pragmatic correction approach based on the z-string-scanning data. Importantly, this simple task with low cognitive demands yielded pupil responses very similar to those of the artificial eye (see Fig. 1a and the Results section above), which allows us to assume that a function fitted on the z-string data can also be a good estimate of the gaze-dependent measurement error.

In the following, we fitted the synthetic correction function (as shown above for the artificial eye) to the z-string-scanning data. This fit resulted in a second- and a fourth-order polynomial function for the left and the right sides of the screen, respectively, which fitted the z-string data nearly perfect ($R^2 = .93$ and $.99$ for the left and the right sides of the screen, respectively; $MSE = 39$ pixels; full

formulas are listed in the Appendix). The present functions will be used subsequently as an alternative estimation of the measurement error for the correction of the theoretically relevant sentence-reading data.

Correction procedure

The synthetic functions, which were described in the previous section, are used to calculate the specific gaze-dependent measurement error for any gaze position during sentence reading. First, the measurement error was estimated by submitting the gaze position of a single measurement to the appropriate function (70–110 pixels from the left screen edge, or = -17° to -15° relative to the screen center, and 110–750 pixels, or -15°–9° to functions 1 and 2, respectively). Second, we subtracted the estimated error from the uncorrected pupil size value. This procedure was realized with the synthetic functions fitted to the artificial eye and the z-string-scanning data.

For illustration, let us assume that the gaze of participant X is at the first word of a sentence and, hence, the actual horizontal gaze position is on the far left side of the screen (e.g., 110 pixels from the left screen edge, or about -15° from the center of the screen). At this particular gaze position, the measurement error would be estimated by the artificial-eye-based synthetic function according to the left part of the screen. For the actual position, the function would estimate a gaze-dependent error of 82.6-pixel (about 6.0%) change relative to the baseline. In a second step, the estimated measurement error would be subtracted from the uncorrected pupil size (78-pixel, or 5.7%, change to baseline), resulting in a feasible 4.6-pixel (0.3%) change of pupil size relative to the baseline.

To estimate the quality of the present correction procedure, we corrected the z-string-scanning data with the z-string-based synthetic correction approach. The resulting deviation from zero of this correction is an estimation of the potential error induced by the correction procedure. As was expected, the resulting corrected pupil response does not show any significant change in pupil size (see Fig. 1c, circle symbols; $MSE = 52$ pixels).

Sentence reading and z-string scanning corrected

In the following, we use the synthetic function to correct the systematic measurement error in pupil size during sentence reading. For the estimation of the measurement error, we use both the artificial eye and the z-string-scanning data. A comparison of the corrected pupil response (Fig. 1c, triangle symbols) with the uncorrected pupil response (Fig. 1a, triangle symbols) reveals that both correction approaches altered the pupil response substantially. During sentence reading, the corrected pupil data

does not show a pupil size change from -17° to -13°. Rather, from -13° to -7°, the pupil starts to dilate and reaches maximum dilation at about -7°. Fig. 1c reveals a slight difference between z-string- and artificial-eye-based correction from -12° to -8° (light gray and dark gray triangles, respectively), which, however, was not statistically significant (multiple *t*-tests, all *t*s < 1).

Importantly, the larger pupil size during sentence reading than during z-string scanning that was reported for the uncorrected data above was also found for the corrected data: A comparison of the corrected pupil response reveals a significantly wider pupil for sentence reading than for the z-string scanning from about -10.5° onward (all *t*s > 2.1, all *p*s < .05).

Finally, uncorrected and corrected pupil sizes are shown relative to the time from the appearance of the sentence onward (Fig. 3). When Fig. 3 is compared with Fig. 1, which shows the measured pupil size in relation to gaze position, the apparent difference is that the initial peak in Fig. 1 is no longer present when the pupil response is related to time. This difference can be explained by the individual reading speed of the participants. To exemplify, a faster reader reaches the gaze position, at which the eye tracker measures the largest pupil size, earlier than does a slower reader. Thus, the marked peak in measured pupil size in relation to gaze position is not present when the measured pupil size is related to the time of the appearance of the sentence.

In accordance with the pupil size in relation to gaze position, the uncorrected pupil response relative to time showed a marked decrease after about 400 ms. The
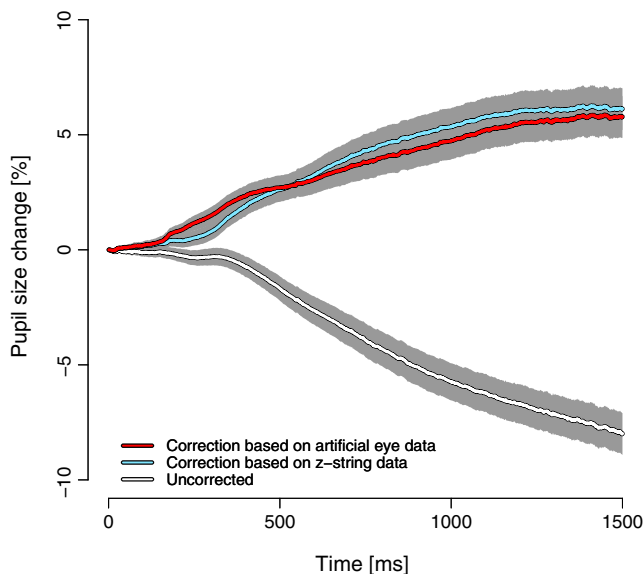

Fig. 3 Means of uncorrected and artificial-eye- and z-string-based corrected pupil responses during sentence reading, relative to reading time (1 *SEM* in light gray)

corrected pupil data, however, indicated a dilatation from around 200 ms until it flattened at around 1,200 ms. There was no significant difference between the two correction procedures (multiple *t*-tests, all *t*s < 1.6). This pattern of results confirms that the correction procedures can also be applied to the pupil response in relation to the time.

In summary, both the artificial-eye-based and the z-string-scanning-based synthetic correction approaches successfully corrected for the gaze-dependent, systematic measurement error in the present study's sentence-reading task. With both correction approaches being of comparable quality, the approach based on z-string scanning (which is easier to realize) can be recommended for future use.

Discussion

The present study investigated the influence of horizontal gaze position on pupil size as measured by video-based eye-tracking systems (i.e., the systematic measurement error) and presented an approach to correct for this systematic measurement error. In general, the present study revealed that the systematic measurement error induced by changes of the horizontal gaze position can be four times the size of the pupil change due to cognitive processing. Therefore, a cognitive interpretation of the uncorrected pupil dilatation data will most likely be misleading.

We quantified the systematic measurement error by measuring the pupil size of an artificial eye (which had a constant pupil size) and, in a more pragmatic approach, by measuring pupil size during an effortless z-string-scanning task. Subsequently, with the intention of providing a correction approach, we successfully modeled the quantified measurement error with a synthetic function. The resulting function was used to correct the pupil size data assessed during sentence reading. The corrected pupil response showed the typical dilation of the pupil, with maximum dilation being reached around 1,200 ms after the presentation of the stimulus. This corrected pupil response corresponds to pupil responses reported by studies that investigated the reading of isolated words (Briesemeister et al., 2009; Kuchinke et al., 2007; Võ et al., 2008).

Up to now, only a few studies that utilized pupil dilation as an index of processing demands used whole sentences (or fragments of sentences) as stimulus material. Most of these studies reported a pupil response that did not show the predominantly reported pupil dilation after the presentation of isolated words. The studies by Raisig et al. (2007) and Schluroff et al. (1986) were singled out here because they were the only studies that reported the pupil response from stimulus onset to offset, which allows a comparison with the findings of the present study. Similar to the uncorrected data in the present study, the authors report an initial

constriction of the pupil during processing of sentence fragments (which had to be read from left to right on the screen) and interpreted this pattern of results as "reflecting the habituation to the situation" (Raisig et al. 2007, p. 869). The findings from the present study (more specifically, the artificial eye data), however, indicate that the initial constriction of the pupil—which is comparable to the pupil responses in Raisig et al. (2007) and Schluroff et al. (1986)—can be attributed to the systematic measurement error that results from changes in gaze position. For future research, the correction functions to overcome this systematic influence are now available, which should facilitate the cognitive interpretation of the pupil responses in tasks that involve eye movements.

When assessments of pupil size are possibly distorted by changes in horizontal gaze position, we strongly recommend exploring the susceptibility of the specific eye-tracking setup that is used for experimentation. This can be realized by an effortless, perceptual task that poses similar demands on visual perception and oculomotor control as the theoretically relevant experimental task, but does not require the cognitive processes under investigation. In the present study, a z-string-scanning task was chosen as the control task for the theoretically relevant sentence-reading paradigm. Other experimental settings might require different control tasks. To illustrate, if one wants to study the pupil response during the free exploration of, for example, scenes or faces, we would recommend exploring the influences of horizontal and vertical changes of gaze position on pupil size. For such an experiment, a simple dot-tracking task might be sufficient. After the exploration of the systematic measurement error, it would be advisable to place the experimental stimuli of the theoretically relevant task in that horizontal and vertical range in which the systematic measurement error can be corrected with a simple set of correction functions. For example, in the present study, it would have been advisable to place the initial word of the sentences at around -14°. Thereby, we could have corrected the systematic measurement error with a single synthetic function, instead of two separate functions. In general, we would recommend the synthetic function (used in correction approach B) for correction, since it provided a superior fit, as compared with the physical function. For the actual correction of pupil size data, the fitted function would be used to predict the measurement error for a specific horizontal gaze position. This predicted measurement error would then be subtracted from the experimental data.

The main focus of the present work was to investigate how horizontal gaze position affects measurement of the pupil size. In the present study, a sentence-reading task was used. As one alternative for correcting artifacts in pupil measurement, we suggested using the z-string-scanning paradigm as an oculomotor control condition for normal reading. We demonstrated that the z-string-scanning data could be used to correct the gaze position artifact for sentence-reading data. Furthermore, the eye movement measures obtained in the z-string-scanning task were in accordance with those in previous studies using this paradigm (Nuthmann et al., 2007; Rayner & Fischer, 1996; Vitu et al., 1995). As a new finding, we reported that the pupil response during z-string scanning resembled the data obtained with the artificial eye and differed from the pupil response during sentence reading insofar as the pupil remained significantly smaller during z-string scanning. These findings support the validity of the mindless reading paradigm as a low-level oculomotor control condition for normal reading (cf. Nuthmann & Engbert, 2009).

To conclude, the present study reveals a substantial, systematic influence of horizontal gaze position on pupil size as measured by video-based eyetrackers during tasks that require eye movements. The study furthermore provides an approach for correcting this influence of gaze position, and the corrected pupil responses will allow more precise cognitive interpretations in future pupillometry studies.

Appendix

Z-string-based synthetic function (70–110 pixels):

$$E_{GPsynth} = -134.2475 + 1.0195*G_P + 0.0083*G_P{}^2 \qquad (9)$$

Artificial-eye-based synthetic function (70–110 pixels):

$$E_{GPsynth} = -172.2500 + 2.0000*G_P + 0.0025*G_P{}^2 \qquad (10)$$

Z-string-based synthetic function (110–850 pixels):

$$E_{GPsynth} = 71.7936 + 0.4283*G_P + 0.0047*G_P{}^2$$
$$+8.2303*10^{-6}*G_P{}^3 - 4.3923*10^{-9}*G_P{}^4 \qquad (11)$$

Artificial-eye-based synthetic function (110–850 pixels):

$$E_{GPsynth} = 185.4830 + 1.0733*G_P + 0.0009*G_P{}^2$$
$$+ 0.0204*10^{-6}*G_P{}^3 \qquad (12)$$

References

Atchison, D., & Smith, G. (2000). *Optics of the human eye*. Oxford: Butterworth-Heinemann.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91,* 276–292.

Briesemeister, B. B., Hofmann, M. J., Tamm, S., Kuchinke, L., Braun, M., & Jacobs, A. M. (2009). The pseudohomophone effect: Evidence for an orthography–phonology-conflict. *Neuroscience Letters, 455,* 124–128.

Granholm, E., Asarnow, R., Sarkin, A., & Dykes, K. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology, 33,* 457–461.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science, 143,* 1190–1192.

Jay, B. S. (1962). The effective pupillary area at varying perimetric angles. *Vision Research, 1,* 418–424.

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indexes of sentence processing. *Canadian Journal of Experimental Psychology, 47,* 310–339.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science, 154,* 1583–1585.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology, 16,* 262–284.

Kuchinke, L., Võ, M. L.-H., Hofmann, M., & Jacobs, A. M. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology, 65,* 132–140.

Nuthmann, A., & Engbert, R. (2009). Mindless reading revisited: An analysis based on the SWIFT model of eye-movement control. *Vision Research, 49,* 322–336.

Nuthmann, A., Engbert, R., & Kliegl, R. (2007). The IOVP effect in mindless reading: Experiment and modeling. *Vision Research, 47,* 990–1002.

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing.*International Journal of Human–Computer. Studies, 59,* 185–198.

Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Psychophysiology, 11,* 559–566.

Piquado, T., Isaacowitz, D., & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology, 47,* 560–569.

Porter, G., Troscianko, T., & Gilchrist, I. D. (2007). Effort during visual search and counting: Insights from pupillometry. *Quarterly Journal of Experimental Psychology, 60,* 211–229.

Raisig, S., Welke, T., Hagendorf, H., & van der Meer, E. (2007). Investigating dimensional organization in scripts using the pupillary response. *Psychophysiology, 44,* 864–873.

Raisig, S., Welke, T., Hagendorf, H., & van der Meer, E. (2010). I spy with my little eye: Detection of temporal violations in event sequences and the pupillary response. *International Journal of Psychophysiology, 76,* 1–8.

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics, 58,* 734–747.

Scheepers, C., & Crocker, M. (2004). Constituent order priming from reading to listening: A visual-world study. In M. Carreiras & C. Clifton Jr. (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP, and beyond* (pp. 167–185). Hove, U. K.: Psychology Press.

Schluroff, M., Zimmermann, T. E., Freeman, R. B., Jr., Hofmeister, K., Lorscheid, T., & Weber, A. (1986). Pupillary responses to syntactic ambiguity of sentences. *Brain and Language, 27,* 322–344.

Stanners, R., Coulter, M., Sweet, A., & Murphy, P. (1979). The pupillary response as anindicator of arousal and cognition. *Motivation and Emotion, 3,* 319–340.

Vitu, F., Oregan, J. K., Inhoff, A. W., & Topolski, R. (1995). Mindless reading: Eye-movement characteristics are similar in scanning letter strings and reading texts. *Perception & Psychophysics, 57,* 352–364.

Võ, M. L.-H., Jacobs, A. M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., et al. (2008). The coupling of emotion and cognition in the eye: Introducing the pupil old/new effect. *Psychophysiology, 45,* 130–140.