Visual sampling processes revisited
replicating and extending senders (1983) using modern eye-tracking equipment

Eisma, Yke Bauke; Cabrall, Christopher D.D.; de Winter, Joost C.F.

DOI
10.1109/THMS.2018.2806200

Publication date
2018

Document Version
Final published version

Published in
IEEE Transactions on Human-Machine Systems

Visual Sampling Processes Revisited: Replicating and Extending Senders (1983) Using Modern Eye-Tracking Equipment

Yke Bauke Eisma*, Christopher D. D. Cabrall*, and Joost C. F. de Winter

***Abstract*—In pioneering work, Senders (1983) tasked five participants to watch a bank of six dials, and found that glance rates and times glanced at dials increase linearly as a function of the frequency bandwidth of the dial's pointer. Senders did not record the angle of the pointers synchronously with eye movements, and so could not assess participants' visual sampling behavior in regard to the pointer state. Because the study of Senders has been influential but never repeated, we replicated and extended it by assessing the relationship between visual sampling and pointer state, using modern eye-tracking equipment. Eye tracking was performed with 86 participants who watched seven 90-second videos, each video showing six dials with moving pointers. Participants had to press the spacebar when any of the six pointers crossed a threshold. Our results showed a close resemblance to Senders' original results. Additionally, we found that participants did not behave in accordance with a periodic sampling model, but rather were conditional samplers, in that the probability of looking at a dial was contingent on pointer angle and velocity. Finally, we found that participants sampled more in agreement with Nyquist sampling when the high bandwidth dials were placed in the middle of the bank rather than at its outer edges. We observed results consistent with the saliency, effort, expectancy, and value model and conclude that human sampling of multidegree of freedom systems should not only be modeled in terms of bandwidth but also in terms of saliency and effort.**

***Index Terms*—Attention, eye tracking, human factors, human–machine systems, multitasking.**

I. Introduction

TECHNOLOGICAL systems are automated to ever greater extents [1]. In many automated systems, the role of the human is to monitor the instruments in order to assess whether the automation performs satisfactorily [2]. Present-day automated systems, such as aircraft cockpit, produce much more information than a human can process at once [3]. Consequently, the human needs to distribute attention across multiple sources of information in order to maintain accurate awareness of the automation state.

How humans sample dynamic instruments is a question that has been of broad interest in human factors and ergonomics (e.g., [4], [5]). Especially in the aviation domain, several studies have been performed that investigated how pilots distribute their visual attention across the different instruments in the cockpit [6]–[9]. In a seminal study, Fitts *et al.* [10] examined how 40 pilots distributed visual attention across cockpit instruments during aircraft landings. Based on their findings, Fitts *et al*. [10] argued that the number of eye fixations per second on an instrument is a measure of the importance of that instrument for carrying out the flight task. Additionally, the fixation duration on the instrument was regarded as an index of the difficulty in reading and interpreting the particular instrument. As pointed out by Landry [11] and Seeberger and Wierwille [12], the results of Fitts *et al*. [10] have been used to redesign the default layout of the cockpit instrument panel in that the instruments most frequently looked at are placed in the middle of the instrument cluster.

Further pioneering work on human sampling behavior of instruments was carried out by Senders [13]. He used the Nyquist–Shannon sampling theorem [14] to predict how frequently a human needs to sample an instrument in order to keep track of its state. This theorem can be intuitively understood when trying to reconstruct a sine wave from a number of periodically sampled data points of this sine wave. If not sampling with at least twice the frequency of the sine wave, then the sine wave cannot be reconstructed from those data points. Accordingly, Senders [15] postulated that if an instrument provides information with a frequency bandwidth W, the human as a Nyquist sampler (ideal observer) should observe that signal with a frequency equal or greater than 2 W.

To test his theory, Senders [13] conducted an experiment in which five undergraduate students monitored a bank of four circular dials (micro-ammeters), with randomly moving pointers that differed in bandwidth (0.08, 0.16, 0.32, and 0.64 Hz). The participants were instructed to press a response key (see Fig. 1) each time one of the four pointers crossed a threshold value from either side. They performed this monitoring task for one hour per day for 30 days. A 3-minute data sample of camera recordings pointed at the eyes of the subjects was collected and analyzed per hour of monitoring. The results

Manuscript received February 28, 2017; revised July 18, 2017 and October 7, 2017; accepted January 14, 2018. Date of publication April 4, 2018; date of current version September 14, 2018. This paper was recommended by Associate editor Prof. Jenna Marquard. *(Corresponding author: J. C. F. de Winter.)*

Y. B. Eisma is with the Faculty of Aerospace Engineering, Delft University of Technology, Delft 2628 CD, The Netherlands (e-mail: y.b.eisma@tudelft.nl).

C. D. D. Cabrall and J. C. F. de Winter are with the Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft 2628 CD, The Netherlands (e-mail: c.d.d.cabrall@tudelft.nl; j.c.f.dewinter@tudelft.nl).

*Y. B. Eisma and C. D. D. Cabrall are co-first authors.

The appendices cited in this paper are available at http://doi.org/10.1109/THMS.2018.2806200. Furthermore, this paper has additional supplementary downloadable material available at http://doi.org/10.4121/uuid:63affb79-d408-4f5b-9b79-8238dd42fa76.

Digital Object Identifier 10.1109/THMS.2018.2806200

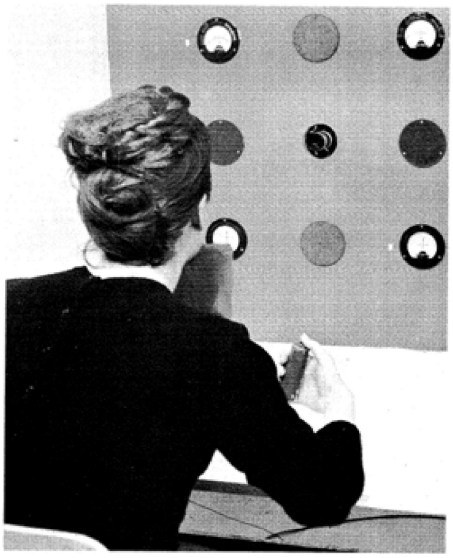


Fig. 1. Illustration of one of the participants in a four-dial sampling task (photo from [17]). A motion picture camera is located in the middle of the four dials. The participant holds a switch in her right hand.

revealed a strong linear relationship between the signal bandwidth (W) and the average observed glance rate (GR) per dial (GR $= 0.05 + 2.44$ W, $r = 0.98$), offering clear support for Senders' theory. Moray [16] suggested that, because eye movements are so strongly predicted by signal bandwidth $(r = 0.98)$, Fitts *et al.* [10] may have been mistaken in that not the importance (e.g., value, cost of missing) of an instrument, but rather its experienced bandwidth (i.e., expectancy) is the prime determinant of how frequently the human looks at an instrument. Put simply, it is possible that pilots in Fitts *et al.* looked at particular instruments more often than at other instruments not necessarily because these instruments were important for the flight task, but rather because these instruments had fast-moving pointers. However, this hypothesis could not be tested because the actual values of the instrument pointers were not recorded by Fitts and co-workers.

In his Ph.D. thesis published almost 20 years later, Senders [15] presented the results of four additional experiments also carried out in the 1960s [17]. These additional experiments were performed using five high school students who viewed six dials of different bandwidths (0.03, 0.05, 0.12, 0.20, 0.32, and 0.48 Hz). These four experiments were similar to each other, but differed somewhat in composition (i.e., a baseline experiment was performed, in a second experiment the random signals were generated in a slightly different manner, in a third experiment a binary signal was used for the 0.12 Hz dial, and in a fourth experiment the bandwidths were slightly varied). Participants received extensive training of at least 10 h. The results of the four aggregated experiments again yielded a nearly perfect linear relationship between bandwidth and glance rate $(r = 0.99)$, but with a slope that was considerably shallower (GR $= 0.18 + 0.61$ W) than predicted by the Nyquist–Shannon theorem and Senders' 1964 experiment (2.44 W). Relative to the model predictions, the shallower slope indicates that participants oversampled the low bandwidth dials while undersampling the high bandwidth dials. One explanation for the undersampling could be that participants tended to forget the state of the low-frequency signals [15], [16]. Furthermore, according to Senders, the introduction of the two very low bandwidth signals (0.03 and 0.05 Hz) may have increased the demands on participants to memorize the state of these dials, in turn causing them to pay less attention to the high bandwidth dials.

Another explanation for a slope shallower than 2 W is the notion that participants may have been able to read the angular velocity of the pointers in addition to the pointers' current angle. This may have reduced the required sampling frequency from 2 W to W [15], [18]. This extension of the sampling theorem can again be intuitively understood by trying to reconstruct a sine wave. If periodically sampling data points of the sine wave, plus the slope of said data points, then the original sine wave can be reconstructed when sampling only once per ordinary frequency of the sine wave. However, the extended sampling theorem cannot explain the different slopes found between Senders' four dial and six dial experiments.

Senders [13], [15] noted that although humans sample in accordance with a periodic sampling model (for the four dial configuration), it is unlikely that humans are actually periodic samplers who deterministically reconstruct a signal according to the sampling theorem, and who do not adjust their sampling behavior based on the momentary state of the pointers. In his thesis, Senders [15] proposed a number of "conditional sampling" models that predict the probability of sampling a particular dial as a function of the current state of the dial relative to the threshold, rather than its overall stochastic property (i.e., bandwidth). Moray [16] eloquently explained why conditional sampling models are viable: "suppose that an observation shows that the function is very close to the permissible limit. It seems likely that another fixation on that source would be made sooner than if it had been observed at, say, its mean" (pp. 40:11). However, because the technology of the 1960s did not allow for a synchronized recording of eye movements and the state of the dials, it still remained to be tested whether conditional models are more valid than a periodic sampling model that uses bandwidth as input. As noted by Senders [15]: "It is necessary to record not only the positions of the eyes but also the value of the signals which are observed. It is only the relationship of these two sets of data that will tell us whether there is anything at all in the idea that observers make use of the information that they see in deciding when to look again." (p. 98).

Various other researchers have proposed conditional models of visual sampling. For example, Carbonell [19] devised a queuing model in which different instruments compete for human attention. The model assumes that each time the human looks at an instrument, he or she postpones the observation of the other instruments, hence accepting the risk that another instrument exceeds a threshold. The optimal sampling strategy is then to sample, and bring back to zero, the instrument with the highest risk of not being observed. The momentary risk per instrument is defined in terms of the cost of exceeding a threshold value (cf., "importance" in [10]) and the probability that the instrument pointer will exceed the threshold, which accumulates as a function of the time since last sampling the instrument.

Carbonell's model was experimentally validated by Carbonell *et al*. [20] but has received little attention since then. Other models of visual sampling behavior were proposed by Sheridan [21] and Kvålseth [22]. However, their models have not been empirically evaluated using eye-tracking equipment.

Nowadays, ample research exists on the topic of visual attention. Borji and Itti [23] provided a review of more than 60 models of visual attention, most of which are bottom-up models (i.e., saliency models). In the last decades, several promising models that include elements of top-down (i.e., task-driven) attention have been developed. For example, Wolfe [24] presented a model that predicts reaction times in tasks where observers look for a target among distractor items. Similarly, Najemnik and Geisler [25] showed that humans can localize a target stimulus embedded in a cluttered environment in an efficient manner, by making eye movements that gain the most information about target location. Salvucci and Taatgen [26] presented a computational model that computes reaction times and performance for diverse multitasking conditions, whereas Sprague *et al.* [27] presented a model of visual behavior, which included a simulated humanoid that allocates gaze based on variables of reward and uncertainty. In an attempt to combine bottom up and top down cues in a comprehensive manner, Wickens *et al*. [28] introduced the saliency, effort, expectancy, and value (SEEV) model of visual behavior. This model defines the probability of sampling an instrument/area in terms of two bottom-up variables: 1) saliency (i.e., the extent to which the stimulus stands out with respect to its background) and 2) effort (i.e., the amount of eye/head movement required) and two top-down variables: 3) expectancy (equivalent to bandwidth, i.e., the perceived likelihood of change or event frequency) and 4) value (i.e., subjective importance of attending to events on the instrument, or the cost of missing them). The SEEV model has received widespread experimental support (e.g., [28]–[30]).

In summary, in the past decades, various models have been developed that describe how humans sample a dynamic system. Much of the current visual models of human monitoring seem to be conceptually based on the original studies by Senders [13], [15] (e.g., [4], [31]). Indeed, the work of Senders is relatively influential in the human factors community, as demonstrated by the ample number of citations in Google Scholar (254 for Senders [13], and 144 for Senders [15]). Perhaps somewhat peculiarly, the work of Senders has hardly been replicated. An exception is Fleetwood [32], who performed three experiments using five participants each. In each experiment, participants viewed four dials as in [13] while eye movements were recorded with a head-mounted eye-tracker. The results of Fleetwood's experiments showed that participants' mean glance durations per dial were sensitive to various experimental manipulations, including bandwidth, threshold cross frequency, value (i.e., different points could be earned based on the correct detection of a pointer that had gone out of bounds), visual saliency (i.e., flashing dial), and the cost of making an observation (i.e., implemented as a time delay when a participant indicated that they would like to view a new dial). Although the design of Fleetwood is regarded as informative, it remains unclear to what extent Senders' results were replicated. Thus, considering that the experiments of Senders were conducted with only five participants and with limited hardware equipment, Senders' study deserves to be replicated and extended for the sake of better insight in human sampling behavior.

The aim of the present study was to replicate the experimental conditions of Senders [15], using a larger sample size ($N = 86$ versus 5) and an eye tracker camera with high temporal resolution (2000 versus 12 Hz). Additionally, whereas Senders' work was solely concerned with coarse dependent measures (i.e., glance rate and duration), we applied fully synchronized data recordings of 1) the six pointer signals, 2) participants' eye movements, and 3) participants' button press inputs. This allowed us to examine how participants distributed their attention across the dials as a function of the state of the dials. An additional factor is that we varied the eye-movement effort level (i.e., one of the parameters in the SEEV model) by changing the dial configuration from a low effort configuration (high bandwidth dials in the center of the bank of dials) to a high effort configuration (high bandwidth dials in the corners of the bank of dials).

In order to structure and interpret our results, we classified our findings according to three variables of the SEEV model: 1) bandwidth (expectancy), 2) effort, and 3) saliency. In our experiment, bandwidth and effort are independent (i.e., experimentally manipulated) variables, whereas saliency is referenced by the momentary state of pointers. Note that value is not an experimental variable in our study, nor in Senders' work: all six dials were assumed to have equal value (i.e., equal importance) for performing the task.

It is noted that part of the results of the same experiment is presented by Eisma *et al*. [33] in more concise form. Therein, Eisma *et al.* [33] were concerned with the broader methodological topic of assessing situation awareness through a correlation between an aggregate visual sampling-to-environmental relational score in comparison to self-reported situation awareness using a freeze-probe questionnaire method. The present study is not concerned with these self-reports but only with objective measures: stimulus behavior, observer performance, and eye-movement data.

II. Methods

A. Participants

Participants were 86 university students (21 female, 65 male) with a mean age of 23.44 years (SD = 1.52). The research was approved by the Human Research Ethics Committee of the TU Delft under the title "Update of Visual Sampling Behavior and Performance with Changing Information Bandwidth" (September 22, 2016). Written informed consent was obtained from all participants.

B. Apparatus and Procedures

The eye movements of the right eye were recorded at 2000 Hz using the SR Research EyeLink 1000 Plus eye tracker. Participants were asked to put their head in a head/chin rest support, which was adjusted to the participant's height to reduce neck

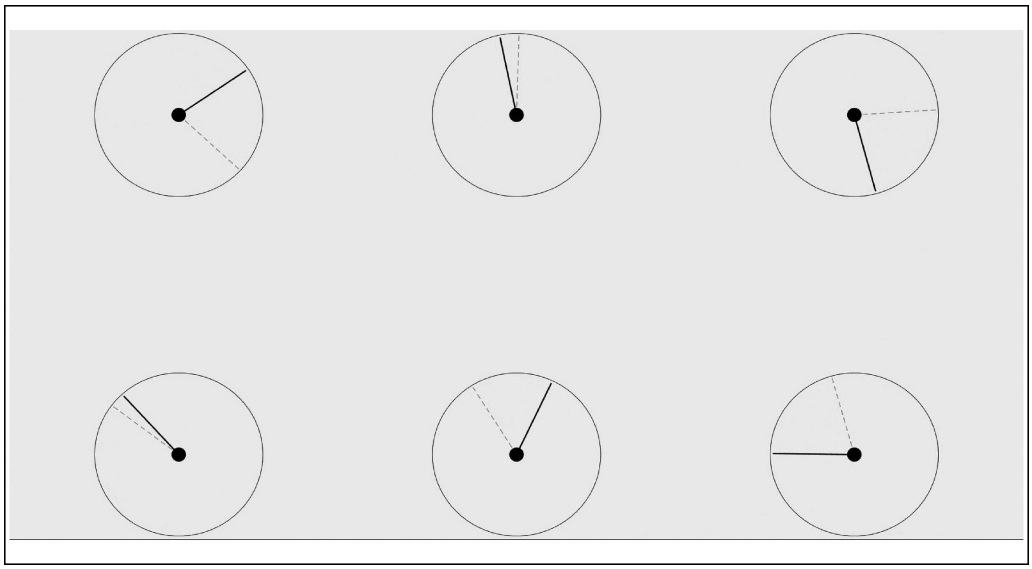


Fig. 2. Screenshot of one of the seven videos. In each dial, the dashed line is the threshold and the solid line is the pointer.

and shoulder strain. The participants were asked to keep their head on the head support throughout the duration of the experiment to the best of their ability, allowing for breaks to counteract any discomfort if needed.

The stimuli were presented on a 24 in BenQ XL2420T-B monitor with a resolution of 1920 × 1080 pixels (display area 531 × 298 mm^2), positioned approximately 95 cm in front of the participant and 35 cm behind the eye-tracking camera/IR light source. The stimulus display subtended approximately a 31° and 18° horizontal and vertical viewing angle, respectively.

First, the eye tracker was calibrated. Next, participants completed a 20 s familiarization trial, allowing them to get used to the experimental setup and task requirements. During this trial, a single dial was shown on the screen.

Next, participants viewed seven 90 s videos. Each video showed six circular dials with moving pointers. Each dial had a diameter of 316 pixels (visual angle ∼5.3°), see Fig. 2. The centers of adjacent dials were 634 pixels (∼10.5°) and 658 pixels (∼10.9°) apart in horizontal and vertical direction, respectively, which is similar to [15] who reported that the dials in his experiments were separated by 12°. The dashed threshold line was a random angle that differed for each of the 42 dials (7 videos × 6 dials). In each of the seven videos, the pointer signals had a mean of 0° (i.e., the position of the threshold) and a standard deviation of 50.1°. The signal realization was different for each of the 42 dials. The MATLAB script that was used for creating the videos is provided in Appendix A.

The frame rate of the videos was 50 Hz, with a resolution of 1904 × 988 pixels. Each participant viewed the same seven videos but in a uniquely randomized order. Participants were instructed to press the spacebar when any of the pointers crossed the threshold from either direction.

After viewing each video, participants completed a brief questionnaire to probe their self-reported situation awareness, knowledge confidence, and experienced eye movement effort. The total time of experiment participation varied between 15 and 30 min.

C. Independent Variables

1) Bandwidth: The first independent variable was the bandwidth of the dials. The six pointers each had a different bandwidth: 0.03, 0.05, 0.12, 0.20, 0.32, and 0.48 Hz, as in [15]. More specifically, a signal was defined as a sum of 21–41 sinusoids, with random phase shifts and with predefined bandwidth (i.e., cutoff frequency) in agreement with Elkind [34] and Senders [15]. Naturally, the high bandwidth dials also moved more rapidly, with overall mean absolute angular pointer velocities of 6.2, 7.3, 13.6, 20.8, 35.4, and 43.5°/s for the 0.03, 0.05, 0.12, 0.20, 0.32, and 0.48 Hz dials, respectively.

The videos are available as the Supplementary material (see Appendix B). The pointer movement of each of the six dials for one of the seven videos is shown in Fig. 3.

2) Effort: The second independent variable was the effort level. Each of the seven videos had a dial configuration that differed according to the predicted amount of eye-movement effort participants had to put in, in order to respond perfectly to each threshold crossing. The configurations were selected with the help of a computer simulation (see Appendix C for the script), in which a value of 1 was assigned to the distance between two adjacent dials (e.g., the diagonal distance between two corner dials was determined as $\sqrt{5}$, see Appendix D for an overview of distances between pairs of dials). All dial configurations are shown in Table I.

Note that Senders similarly positioned the dials "in a quasi-random way in order to achieve as much counterbalancing as possible, since the theoretical model, which was to be tested, did not consider the factor of *arrangement* of signals of various frequencies" (see [17, p. 44]). However, Senders [15], [17] did not present the actual dial configurations.

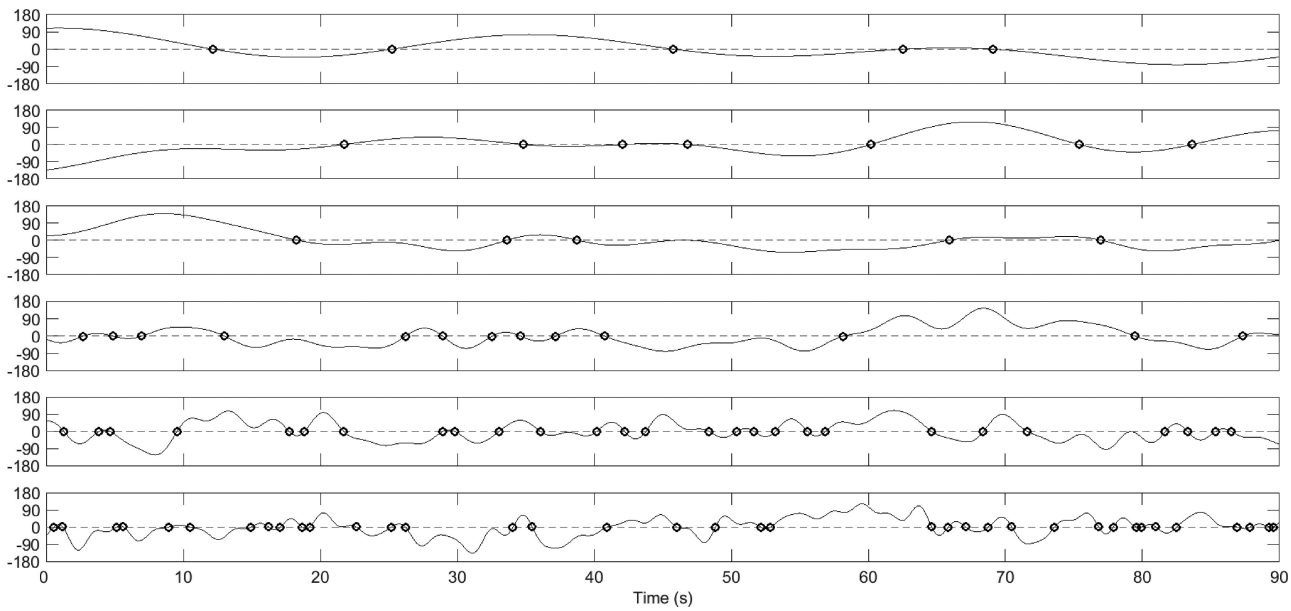


Fig. 3. Pointer angle in degrees relative to the threshold (positive = clockwise with respect to the threshold, negative = counterclockwise with respect to the threshold) as a function of elapsed time in one of the videos. The six subplots are sorted on bandwidth (top = 0.03 Hz, bottom = 0.48 Hz). A circular marker indicates a threshold crossing.

TABLE I
BANDWIDTH (HZ) PER DIAL POSITION FOR EACH OF THE SEVEN VIDEOS USED IN THE EXPERIMENT

Video effort level	Top left	Top middle	Top right	Bottom left	Bottom middle	Bottom right	Effort level
Level 1 (lowest effort)	0.12	0.48	0.05	0.20	0.32	0.03	3422
Level 2	0.20	0.48	0.03	0.32	0.05	0.12	3686
Level 3	0.03	0.12	0.20	0.32	0.48	0.05	3896
Level 4	0.32	0.12	0.05	0.48	0.03	0.20	4097
Level 5	0.48	0.05	0.03	0.12	0.20	0.32	4314
Level 6	0.12	0.32	0.05	0.20	0.03	0.48	4532
Level 7 (highest effort)	0.32	0.03	0.20	0.12	0.05	0.48	4969

The video numbers range from low effort (high bandwidth dials in the middle) to high effort (high bandwidth dials in the outer edges). The effort level is the estimated cumulative saccade distance if the participant were to sample perfectly for 1 h of observation.

D. Dependent Variables

1) Dependent Measures to Replicate Senders [15]: First, missing x and y coordinates during blinks were restored with linear interpolation. Furthermore, a median filter with a 100 ms interval was applied to the x and y gaze coordinates. Next, the following measures were calculated per participant, per dial, and per video:

1) *Glance rate (Hz)*, defined as the number of times per second that the participant fixated on a 420 × 420 pixel area of-interest (AOI) surrounding the dial. Refixations on the same dial were not counted. By virtue of a fixation filter, only glances on dials were counted, not fly-throughs (e.g., the top middle dial was not counted when the participant performed a saccade from the top left to the top right dial). Gaze velocity data were calculated and filtered with a Savitzky–Golay filter with order 2 and a frame size of 20 ms (i.e., 41 samples at 2000 Hz, twice the minimum saccade duration of 10 ms, see [35]). We adopted a saccade velocity threshold of 2000 pixels/s (∼33°/s). It has been reported that fixation durations in reading can be as short as 50–75 ms [36]. Considering that the present task involved rapid sampling and small visual angles, a minimum fixation duration of 40 ms was used, see also [35].
2) *Percent time on AOI (%)*, defined as the percentage of video time that the eye-gaze of the participant was within a specified dial AOI. This measure was calculated independently from the fixation filter and has also been referred to as the *net* dwell time percentage [37].
3) *Mean glance duration (s)*, defined as the net dwell time per dial in seconds divided by the number of glances on that dial.

These three preceding measures were compared to the corresponding measures reported in [15]. Note that Senders [15] used the terms 1) fixation frequency or sampling frequency, 2) percent time fixated, and 3) duration of fixation, for the three above-mentioned measures, respectively. However, for the sake of clarity, we adhered to modern terminology in line with standards [38].

Additional dependent measures were taken as follows.

2) Spacebar Press Performance: We calculated a performance score, defined as the percentage of threshold crossings for which the participant pressed the spacebar. In total, there were between 74 and 115 threshold crossings per video. Per crossing, a hit was counted if the participant pressed the spacebar within 0.5 s (i.e., between −0.5 and +0.5 s) of the moment of the crossing. Specifically, hits were determined using a for-loop over the threshold crossings of a video in a chronological order. For each threshold crossing, the temporally closest spacebar was selected, and if the absolute time difference between the moment of pressing the spacebar and the moment of the threshold crossing was smaller than 0.5 s, then that threshold crossing was labeled a hit, and the spacebar press was excluded from being assigned to subsequent threshold crossings. Accordingly, a spacebar press could not be assigned to more than one threshold crossing, and no more than one hit could be assigned to a threshold crossing.

3) Questionnaire Data: Per participant and per video, we calculated the experienced eye-movement effort. This measure was defined as the response to the question "How much eye-movement effort did you experience?," with response options from 1 (very low) to 10 (very high).

E. Analyses

In order to structure our findings, our analyses were categorized into 1) bandwidth (expectancy), 2) effort, and 3) saliency, which are the first three predictor variables of the SEEV model.

1) Bandwidth (Expectancy)—Replication of Senders [15]: First, we reported the overall glance rate, percent time on AOI, and mean glance duration as a function of bandwidth, in order to examine whether the results of Senders were replicated in our study. Similarities between our results and Senders' results were assessed by comparing the parameters of linear least squares fits between the bandwidth and the dependent measure.

A periodic sampling model assumes that the human observer forms expectancies about the likelihood that a pointer will cross a threshold, or as pointed out by Senders [15]: "in order to make a rational allocation of visual attention to various signals, the observer must learn the bandwidths of those signals" (p. 86). To investigate whether participants exhibited learning (i.e., whether they formed expectancies) during the experiment, we assessed linear least squares fits between glance rate and bandwidth, per video presentation number in the chronological order. This allowed us to assess whether participants distributed their attention more akin to the Nyquist theorem as they gained experience at the sampling task.

2) Effort (Dial Configuration): Similar analyses were conducted for the different video effort levels. That is, we calculated linear fits between the mean glance rate and dial bandwidth, for each dial configuration condition shown in Table I. Additionally, the performance score and self-reported effort were computed per effort level, to see whether participants performed more poorly in the high effort videos than in the low effort videos.

3) Saliency (Pointer Angle, Pointer Velocity, Time to Crossing): Finally, we assessed whether participants were conditional samplers by calculating for each video frame the percentage of participants who glanced at each dial and comparing this to the angle and velocity of the dial pointers at that video frame. Here, pointer angle (i.e., closeness to threshold) and pointer velocity are regarded as components of saliency, that is, the extent to which the current state of a pointer attracts visual attention. Note that for a single sine function, position and its derivative are directly related, as the derivative of a sine wave equals the same sine wave with a phase shift, and hence, in this case, it would be meaningless to analyze the effects of pointer angle and pointer velocity separately. However, in our experiment we used a multisine consisting of 40 aggregated sine waves, as a result of which pointer angle and pointer velocity were not directly related, except at its extreme values (i.e., when a pointer signal reaches its peak angle in a given video, the pointer has a velocity of zero by definition).

In addition to the pointer angle and pointer velocity, we assessed conditional sampling for the "time to crossing," defined as the momentary pointer angle divided by the sign-inverted pointer velocity (see [39] for a similar time to line crossing measure in car driving, and [40] for the notion that humans may be able to perceive time to crossing directly from the closure rate of the log-transformed angle between pointer and threshold). A positive time to crossing means that the dial is moving in the direction of the threshold, whereas a negative value means that the pointer is moving away from the threshold.

It is noted that vision researchers typically use the word saliency to refer to stimulus characteristics such as intensity contrast, flicker contrast, and motion contrast, devoid of task context [41]. Absolute pointer velocity is a saliency feature, but the pointer angle is not. That is, participants should interpret the pointer angle in relation to the task of pressing the spacebar when it crosses the threshold. Herein, we use the term saliency in a broad meaning, by defining it as the dial's momentary characteristics (as opposed to bandwidth, which is a time-invariant property).

III. Results

Data were lost for a few videos (1–3) from a total of three participants. However, because the majority of their data were still available and unaffected, these three participants were retained in the analysis.

A. Descriptive Statistics: Aggregate Gaze Results

Fig. 4 shows the aggregated distribution of all gaze coordinates on the monitor. It is apparent that not all six dials exhibited the same percent time on AOI. The percent time on AOI was the highest for the top middle dial (20.18%) position and the lowest for the top right dial (8.50%) position. These differences are consistent with a baseline tendency to look at the middle two dials, and can also be explained by the different bandwidth configurations per dial (see Table I). For example, the top right dial position never happened to display a high bandwidth dial signal (0.32 or 0.48 Hz), which may explain why it was overall less sampled relative to the dial positions in the other corners. Further analysis (see Appendix D) showed that diagonal eye movements

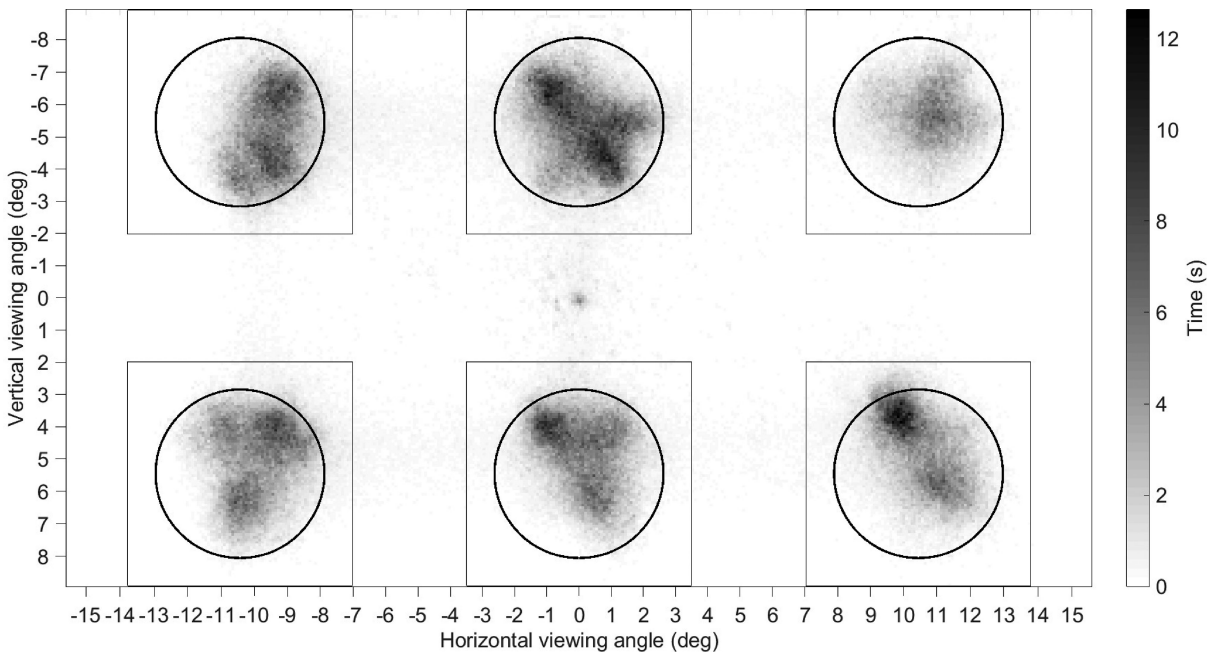


Fig. 4. Distribution of gaze for all videos of all 86 participants aggregated (53 550 s of data). For the purposes of this visualization, the screen was divided into 5 × 5 pixel squares, and the darker the color, the more time was spent looking at that part of the screen as indicated by the vertical bar on the right. The circles represent the dials; the squares that surround the circles represent the areas of interest.

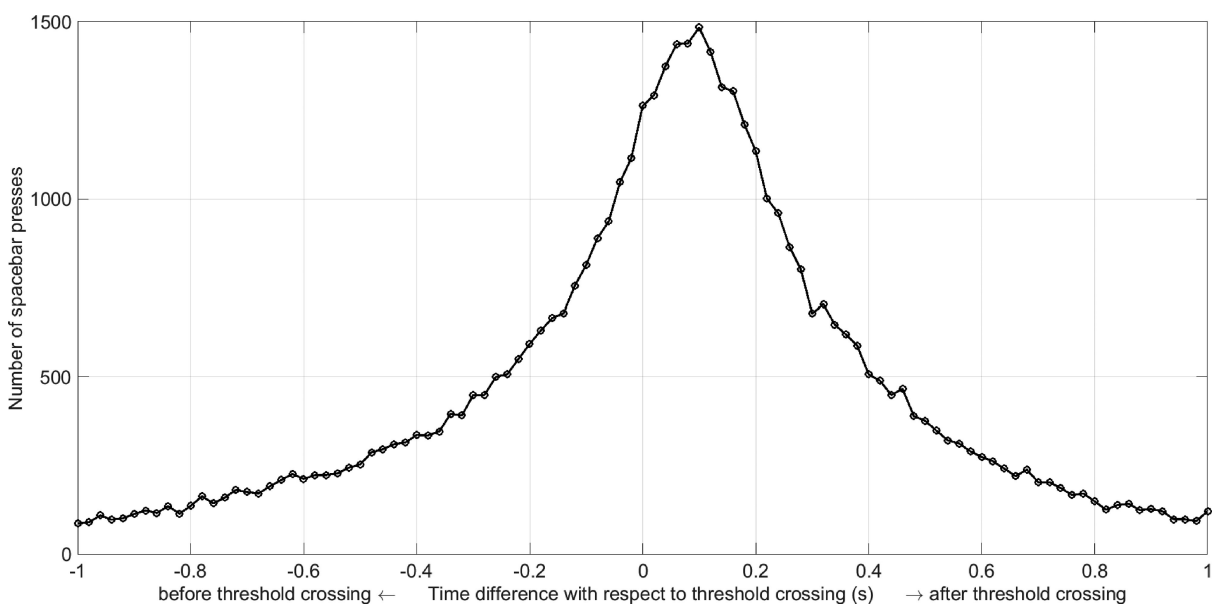


Fig. 5. Time difference between threshold crossing and the spacebar press that occurred nearest in time, for each threshold crossing ($N = 52\,627$). Results are presented in 0.02 s intervals.

were rarer compared to horizontal and vertical ones, see [4] for a similar finding.

B. Descriptive Statistics: Aggregate Performance Results (Spacebar Presses)

Fig. 5 shows the distribution of the time difference between the spacebar presses and the threshold crossings of the pointers. The 10th, 50th, and 90th percentiles of the time difference were −0.66, 0.06, and 0.68 s, respectively. Accordingly, our definition of performance, which incorporated a time margin from −0.5 to 0.5 s surrounding each threshold crossing, is regarded as reasonable in that it captured the majority of spacebar presses surrounding a threshold crossing while minimizing overlap between consecutive threshold crossings.

Table II shows that participants slightly improved their spacebar-pressing performance score from 47.53% during the

TABLE II
PERFORMANCE SCORES AS A FUNCTION OF VIDEO PRESENTATION ORDER (I.E., LEARNING EFFECT IN PERFORMANCE) AND AS A FUNCTION OF THE VIDEO EFFORT LEVEL (I.E., EFFECT OF DIAL CONFIGURATION ON PERFORMANCE)

Performance as a function of video presentation order		Performance as a function of video effort level	
Video order	Performance score (%) M (SD)	Video effort level	Performance score (%) M (SD)
First	47.53 (8.74)	Level 1 (lowest effort)	52.92 (8.39)
Second	48.49 (8.54)	Level 2	51.35 (8.66)
Third	49.40 (9.14)	Level 3	47.97 (8.22)
Fourth	48.42 (9.08)	Level 4	48.49 (8.43)
Fifth	49.82 (8.11)	Level 5	47.28 (8.27)
Sixth	51.40 (8.38)	Level 6	50.01 (8.68)
Seventh (last)	51.17 (8.63)	Level 7 (highest effort)	48.11 (9.23)

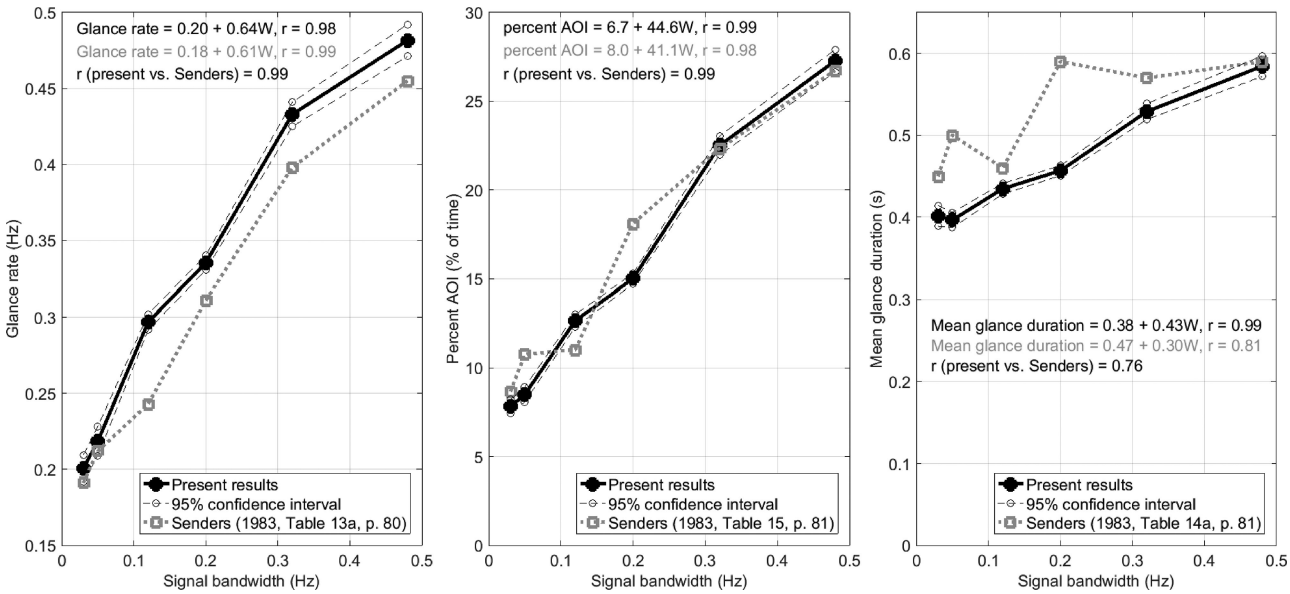


Fig. 6. Glance rate, percentage of time on area of interest (AOI), and mean glance duration as a function of signal bandwidth of the dial. The dashed lines with open circles represent 95% confidence intervals around the mean, calculated according to Morey [41]. The grey dotted line with square markers corresponds to Senders' summary results in which he averaged the results of three similar experiments. The equations represent a least squares linear fit for our results (in black) and Senders' [15] results (in grey). Also shown is the Pearson correlation coefficient between our results and Senders' results.

first video up to 51.17% in the last video, whereas the corresponding standard deviation among participants remained approximately constant. The effect of video presentation order was small but statistically significant according to a repeated measures ANOVA for the 83 participants without missing values, $F(6,492) = 5.37, p < 0.001, \eta_p^2 = 0.061$. Additionally, a higher effort configuration of the dials corresponds with a slightly lower performance score (see Table II). The effect of the video effort level was statistically significant as well, $F(6,492) = 14.14, p < 0.001, \eta_p^2 = 0.147$.

We observed no interpretable relationship between dial position (i.e., top left, top middle, top right, bottom left, bottom middle, bottom right) and performance scores (see Appendix E). However, the lowest bandwidth dial (0.03 Hz) featured a lower performance score (30.96%) than the five higher bandwidth dials (46.55% and higher, see Appendix E). Further inspection revealed that the difficulty of the dials was highly idiosyncratic: among the 42 dials (7 videos $\times$ 6 dials) the performance score ranged between 16.00% (SD $= 16.85\%$) for the top middle dial (0.03 Hz bandwidth) of the video with effort level 7, and 66.09% (SD $= 20.20\%$) for the bottom right dial (0.12 Hz bandwidth) of the video with effort level 2.

C. Bandwidth (Expectancy)—Replication of Senders [15]

Fig. 6 shows the glance rate, percent time on AOI, and mean glance duration, as a function of bandwidth, together with 95% confidence intervals for the means across the participants. Also shown are the results of Senders, which are based on a total of five participants. Because Senders used a small number of participants, no confidence intervals were calculated for his dataset. The results reveal a high correspondence between our results and Senders' [15] results ($r = 0.99, 0.99$, and 0.76 for the three respective measures).

In order to assess whether participants exhibited learning (i.e., whether they formed expectancies of bandwidth) from the first video to the seventh video, the glance rate as a function of bandwidth was assessed per video number. The results in Table III

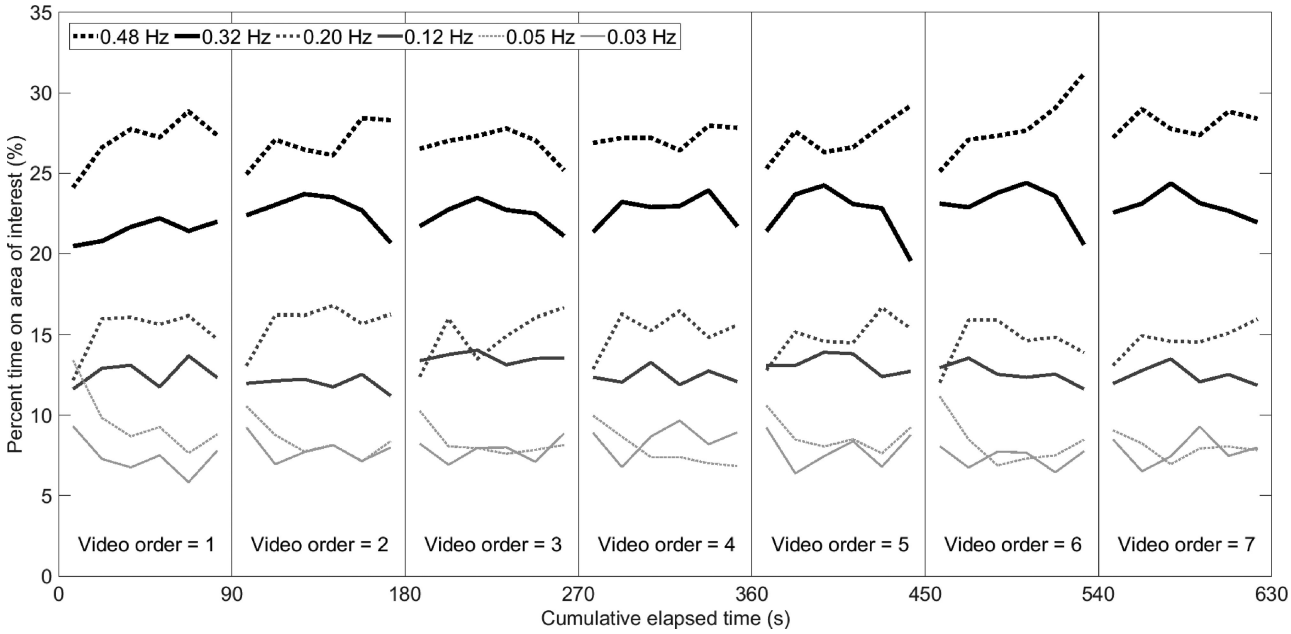


Fig. 7. Percentage of time that participants had their eyes on a particular bandwidth dial as a function of the total elapsed video time. Each video lasted 90 s. The results are provided as averages per 15 s wide bin. These confidence intervals are depicted only for the lowest and highest bandwidth dials, in order to prevent clutter.

TABLE III
LINEAR FIT FOR BANDWIDTH (*W*) VERSUS MEAN GLANCE RATE (GR) AS A FUNCTION OF THE CHRONOLOGICAL ORDER OF VIDEO PRESENTATION

Video presentation order	Linear fit and correlation coefficient (r)
First	$\mathrm{GR} = 0.21 + 0.61\ \mathrm{W}, r = 0.98$
Second	$\mathrm{GR} = 0.20 + 0.64\ \mathrm{W}, r = 0.98$
Third	$\mathrm{GR} = 0.20 + 0.63\ \mathrm{W}, r = 0.98$
Fourth	$\mathrm{GR} = 0.20 + 0.65\ \mathrm{W}, r = 0.97$
Fifth	$\mathrm{GR} = 0.21 + 0.62\ \mathrm{W}, r = 0.98$
Sixth	$\mathrm{GR} = 0.20 + 0.66\ \mathrm{W}, r = 0.98$
Seventh (last)	$\mathrm{GR} = 0.19 + 0.68\ \mathrm{W}, r = 0.98$

TABLE IV
LINEAR FIT FOR BANDWIDTH (*W*) VERSUS MEAN GLANCE RATE (GR), AND SELF-REPORTED EFFORT, AS A FUNCTION OF THE VIDEO EFFORT LEVEL (SEE TABLE I FOR DEFINITION)

Video effort level Video effort level	Linear fit and correlation coefficient (r)	Self-reported effort *M* (SD)
Level 1 (lowest effort)	$\mathrm{GR} = 0.11 + 1.00\ \mathrm{W}, r = 0.97$	6.89 (1.61)
Level 2	$\mathrm{GR} = 0.16 + 0.84\ \mathrm{W}, r = 0.95$	6.84 (1.59)
Level 3	$\mathrm{GR} = 0.21 + 0.65\ \mathrm{W}, r = 0.86$	7.01 (1.62)
Level 4	$\mathrm{GR} = 0.23 + 0.50\ \mathrm{W}, r = 0.82$	6.91 (1.86)
Level 5	$\mathrm{GR} = 0.26 + 0.38\ \mathrm{W}, r = 0.71$	7.16 (1.49)
Level 6	$\mathrm{GR} = 0.18 + 0.66\ \mathrm{W}, r = 0.84$	7.24 (1.62)
Level 7 (highest effort)	$\mathrm{GR} = 0.23 + 0.44\ \mathrm{W}, r = 0.96$	7.19 (1.54)

show that there is a slight learning effect, as the slope is shallowest 0.61 W for the first video and steepest 0.68 W for the seventh video. Note that these changes in the parameters of the linear fits are overall small and that the parameters for all video presentation orders are in agreement with Senders who reported $\mathrm{GR} = 0.18 + 0.61\ \mathrm{W}, r = 0.99$ (see Fig. 6).

Fig. 7 presents a further illustration of the learning effect within a particular video. It can be seen that the percent time on the dials with different bandwidths can already be differentiated from the beginning (i.e., in the first 15 s of each 90 s). There also appears to be a slight periodicity for each of the seven videos as it seems to take about 30 s for AOI percentages to settle in (e.g., the 0.20 Hz dial appears to be relatively undersampled in the first 15 s of each video). A paired t-test between the first 15 s and last 15 s indicated the following: $t(82) = 1.42, 6.03, -0.36, -3.30, -0.90, -2.64$ ($p = 0.160, < 0.001, 0.718, 0.001, 0.370, 0.010$, Cohen's $d_z = 0.16, 0.66, -0.04, -0.36, -0.10, -0.29$) for the 0.03, 0.05, 0.12, 0.20, 0.32, and 0.48 Hz dials, respectively. In other words, the low bandwidth dials tended to be sampled *less* while the high bandwidth dials tend to be sampled *more* in the last 15 s as compared to the first 15 s. Thus, slight learning/habituation effects are distinguishable: sampling becomes more distributed with experience, which is in line with the increasing slope from 0.61 W to 0.68 W shown in Table III.

D. Effort (Dial Configuration)

The results in Table IV show that the slope of the regression line between bandwidth and glance rate was steepest (1.00 W) for the lowest effort configuration and considerably shallower (0.44 W) for the highest effort configuration. To illustrate, in the lowest effort configuration, participants had a glance rate of 0.128 and 0.554 Hz to the low and high bandwidth dial, respectively. In the highest effort configuration, this was 0.252 and 0.429 Hz, respectively. In other words, when the effort was lower, the effect of bandwidth on distributed sampling was

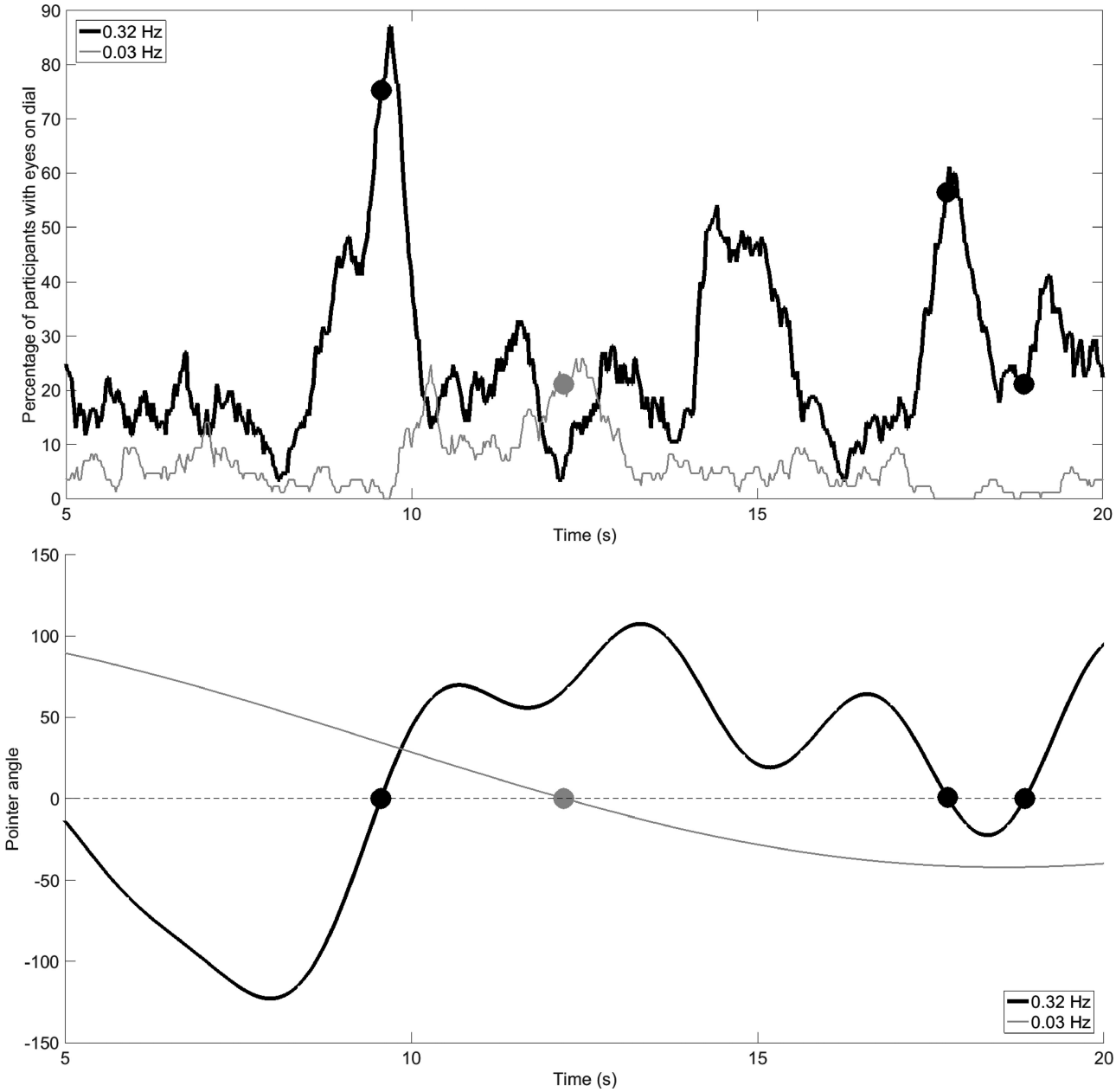


Fig. 8. Percentage of participants ($N = 85$) with their eyes on a dial (top) and state of the dial relative to the threshold (bottom) for a representative 15 s segment of the first of seven videos. The circular markers are depicted at the moments of the threshold crossings.

higher. Thus, when the high bandwidth dials were placed in the middle (e.g., video effort level 1) instead of at the outer edges (e.g., video effort level 7), participants behaved more in accordance with the Nyquist theorem (i.e., a slope that is closer to the theoretically predicted slope of 1 W or 2 W, depending on whether or not sampling of the velocity is taken into consideration). Conversely, when the high bandwidth dials were placed at the outer edges, participants relatively rarely sampled these high bandwidth dials while relatively often sampling the low bandwidth dials in the middle, in line with the notion that effort inhibits sampling.

Table IV also shows that objective effort had a small effect on subjective effort; this effect was significant according to a repeated measures ANOVA, $F(6,498) = 2.78, p = 0.012, \eta_p^2 = 0.032$.

E. Saliency (Pointer Angle, Pointer Velocity, Time to Crossing)

An initial exploration confirmed that participants' sampling behavior was indeed not only dependent on bandwidth and effort, but also highly time-varying. Fig. 8 shows the percentage of participants who gazed at two specific dials for a random 15 s segment of one of the seven videos. Once again, it is evident that participants looked more at a high bandwidth dial than at a low bandwidth dial. Closer inspection shows that participants were more likely to gaze at a particular dial (see peaks in the

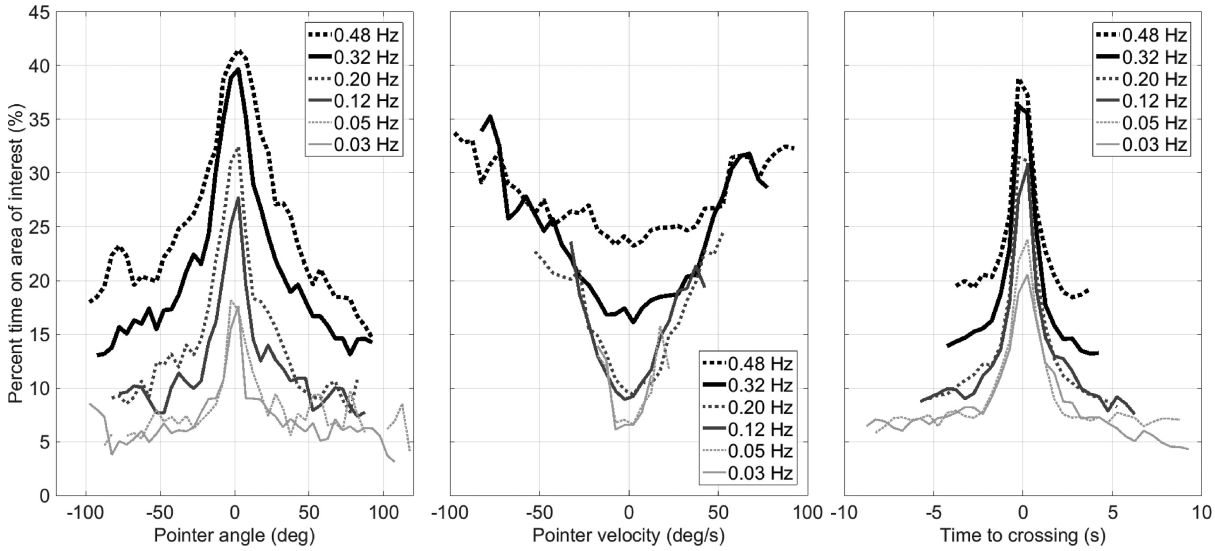


Fig. 9. Relationship between pointer angle in 5° increments (left), pointer velocity in 5°/s increments (middle), and time to crossing in 0.5 s increments (right) versus percent time on area of interest. The results in this figure were based on all videos of all participants. Only data points for which at least 5 s of video data were available are shown.

upper graph) when the pointer angle was near the threshold (see bottom graph having the same time axis as the upper graph). Furthermore, peaks in the upper graph appear to occur when a pointer angle has a high gradient, that is, when the pointer was moving rapidly.

The relationship between the participants' viewing behavior and the state of the dials is further illustrated in Fig. 9. The left panel shows the proportion of participants sampling a specific dial as a function of the pointer angle with respect to the threshold. It is clear that participants were considerably more likely to gaze at a dial when the dial was close to the threshold. When the dial was near the threshold at 0°, the probability of sampling the dial was about 2–3 times as high as compared to when the dial was at an angle of 45° away from the threshold. In the middle and right panel of Fig. 9, this effect can also be seen for the angular velocity of the pointer and time to crossing, respectively. Note that there appears to be no asymmetry: high-velocity pointers attract attention regardless of whether the pointer is moving toward or away from the threshold. Fig. 10 further illustrates that a combination of low pointer angle and high pointer velocity is an attractor of attention. Also, it is notable that the low bandwidth dials never reach a high velocity in the first place.

In sum, participants do not behave in accordance with a periodic sampling model. Rather, participants sample conditionally: the closer the pointer to the threshold, and/or the faster it moves especially toward that threshold, the more the participants gazed at that specific dial.

IV. Discussion

A. Bandwidth (Expectancy)—Replication of Senders [15]

The aim of this research was to replicate Senders' [15] study of visual sampling, using high-end eye-tracking equipment, and a larger number of participants. The results of our experiment showed that the glance rate, the percent time on AOI, and the mean glance duration increase as the bandwidth increases, in close similarity to what was found by Senders (i.e., highly similar slopes and intercepts, and strong correlations between our results, $r = 0.99, 0.99$, and 0.76). In his work, Senders [15] noted that the high bandwidth dials do have a longer duration of observation, but he also expressed considerable uncertainty about this effect. Our results confirm for the first time a linear relationship between bandwidth and mean glance duration ($r = 0.99$, see Fig. 6). Presumably, we obtained a stronger correlation with bandwidth than the correlation obtained by Senders ($r = 0.81$) because Senders used only five participants in his experiment, hence giving rise to a considerable sampling error. Furthermore, Senders used manual coding of film recordings in lieu of eye tracking. These film images were recorded at 12 Hz, which means that the temporal resolution of his method was at best 0.083 s, or perhaps only 0.167 s if considering that at least two frames are needed to ascertain whether the eyes of the participant have actually landed on the dial. In comparison, we used an eye tracker with 2000 Hz resolution, combined with a fully automated data analysis procedure, which is insensitive to manual coding errors.

Senders argued that people need extensive training in order to learn the statistical characteristics (i.e., bandwidths) of the pointers: "The theory and the attendant models, therefore, apply only to demanding tasks performed by experienced and skilled human beings. No novices need apply." (see [15, p. 21], and also [16]). The participants in our experiment were only allowed to familiarize for 20 s with a single dial (as an example), yet our results show considerable similarities with Senders' results. This suggests that participants do not need to learn the statistical characteristics of the signal, but predominantly rely on the momentary state of the dial in order to perform the sampling

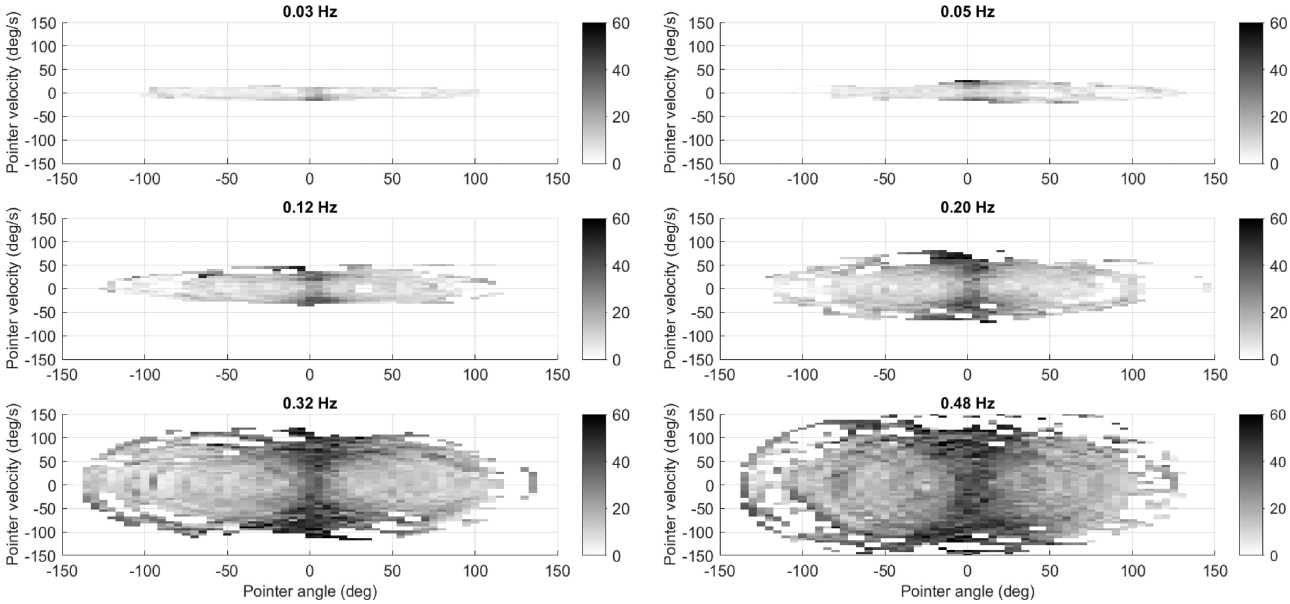


Fig. 10. Percent time on area of interest (as indicated by the vertical bar next to each figure) as a function of pointer position and pointer velocity. The present figure shows the probability that participants sampled a dial for a combination of pointer position and pointer velocity. Pointer angle and pointer velocity were divided into 5° and 5°/s increments, respectively.

task. In our experiment, participants did show learning, as they distributed their attention more according to bandwidth in later sessions (see Table III; Fig. 7). However, this learning effect was minor compared to the effects of bandwidth and effort (dial configuration; Table IV). Furthermore, a comparison between the first 15 s and last 15 s (see Fig. 7) showed that the learning effect was bandwidth-specific (e.g., the 0.05 Hz dial showed a larger learning effect than the 0.03 Hz dial), which may be due to interactions with the dial configuration or the specific properties of the pointer signals.

B. Effort (Dial Configuration)

The information access effort in the context of this experiment can be defined as the amount of eye-movement required in order to detect the events (i.e., the threshold crossings). According to Wickens [42], people tend to minimize effort during the task, hence try "to avoid longer scans or other information access travels when shorter ones can be made" (p. 54). Our results provide support to the notion that objective effort inhibits sampling: people are more likely to gaze at high bandwidth dials when these are placed centrally and generally sample less in accordance to bandwidth as required eye-movement distances grow (see Table IV for the corresponding linear regression results).

C. Saliency (Pointer Angle, Pointer Velocity, Time to Crossing)

Senders [13] originally modeled human sampling behavior by assuming that the human acts as "a random sampling device constrained *only* by the base probabilities of each of the things sampled" (p. 5, emphasis added). However, for a six dial configuration, the slope of glance rate versus bandwidth is considerably shallower than the theoretically predicted 2.0. Both Senders [15] and ourselves found a slope of about 0.60 combined with an intercept of about 0.20 (see Fig. 6), indicating that participants undersampled the higher bandwidth dials and oversampled the lower bandwidth dials relative to an assumed perfectly matched bandwidth dependent sampling behavior (i.e., slope of 2.0 and intercept of 0).

According to Senders [15], this shallow slope may be attributed to 1) mental overload, 2) the fact that participants exhibit forgetting of the pointer state since the last glance on the dial, and also 3) the fact that participants may sample not only the pointer angle but also the pointer velocity (i.e., additional information in the task such as stimulus saliency). The latter explanation is in agreement with the extended sampling theorem [18], which postulates that the slope of an observer is 1 W instead of 2 W when the observer extracts both momentary velocity and momentary position. Our results in Figs. 8–10 indicate that participants were more likely to glance toward a dial when the velocity of that dial's pointer was higher.

In sum, the periodic sampling model is contentious because the probability of sampling is strongly dependent on how close the dial is to the threshold and how rapidly the pointer is moving. It is striking that there is a strong U-shape for low bandwidth dials in particular: for relatively high pointer velocities (around −40 or 40°/s), the dwell percentages for low bandwidth dials are about equal to the dwell percentages for high bandwidth dials (see Fig. 9, middle). In other words, the pointer velocity can be a strong attention attractor even when the dial bandwidth is relatively low.

We argue that participants were able to detect whether something is happening quickly in the periphery, resulting in a state of uncertainty, which in turn attracts attention. The notion of motion being an attention attractor corresponds to the saliency cue in Wickens' SEEV model of visual sampling and many other types of bottom-up visual attention models [23]. Previous research shows that humans can perform a control task such as car driving [43] or pitch tracking [44] using peripheral vision.

Senders *et al*. [45] specifically examined whether it is possible to read a dial using peripheral vision, and one of their conclusions was that "an observer can discriminate among settings which differ by 45° almost perfectly even when the instrument is played as much as 40° from the line of sight" (p. 436). In a pilot experiment using larger dials and red thresholds that were placed upright, we found that a participant could complete the spacebar-pressing task satisfactorily while looking only at the center of the screen. Although the present layout (i.e., smaller dials, dashed threshold at various angles; see Fig. 2) is more difficult to perform using just peripheral vision, it is likely that participants are still able to extract some information from their periphery. In sum, we argue that participants do not have to rely only on the learned bandwidth of a signal to determine where to look (as predicted by the periodic model); rather, they detect in their periphery salient aspects of whether a dial's threshold is likely to be crossed (e.g., from a pointers' velocity, threshold proximity, or closure rate), which in turn attracts their foveal attention toward that dial.

D. Conclusion and Recommendations

Collectively, our results offer a more fine-grained picture of human visual sampling than that of periodic signal reconstruction according to the (extended) Nyquist–Sampling sampling theorem. In particular, our results indicate that even for a simple paradigm of six moving dials, human visual sampling should not be explained in terms only of bandwidth (expectancy) but also by effort and saliency, as used in the SEEV model [28]. In conclusion

1) the results of Senders [15] have been replicated using high-end eye-tracking equipment,
2) humans do not behave as periodic samplers, but as conditional samplers instead, and
3) the conditions upon which humans sample include aspects of both "saliency" and "effort" in addition to "expectancy" (i.e., base bandwidths) when considering visual sampling behavior in goal-directed task environment of a certain performance value.

Future research could be directed toward resolving some uncertainties in the present findings. First, although there is a close correspondence between the results obtained by Senders [15] and the results presented herein, it cannot be established what exactly caused the similarity of results. We closely reproduced Senders' signal composition and task instructions, but there are also some evident differences between our experiments. That is, we used a computer screen, whereas Senders used microammeters, and we used a single randomly oriented threshold per dial (see Fig. 2), whereas Senders used a fixed threshold at about 56° on either side for all six dials.

Additionally, Senders provided participants with more than 10 h of training, whereas we provided essentially no training. In future research, these effects could be studied independently in more detail. It may be worthwhile to investigate how with elongated practice the components of the SEEV model come into play in a different manner. With extended exposure, the impact of saliency might be expected to slightly decrease due to habitation, while the impact of effort may also slightly decrease, as practice may encourage the development of more efficient motor/behavioral patterns to some limit.

The present study suggests that participants sample conditionally rather than periodically, and that peripheral saliency is an important attractor of attention. Future research could examine whether participants still learn some of the signal properties so that they can direct attention as a function of bandwidth even if peripheral vision is unavailable. For example, future research may use occlusion techniques [cf., 46] or a gaze-contingency paradigm.

We also recommend research into different dimensions of effort, such as eye movement effort, head movement effort, and cognitive effort. For example, it would be interesting to examine what happens if head movement is not restricted by a head support. Here, it might be expected that participants will orient their head toward the high bandwidth stimuli, and accordingly mitigate their required sampling effort. Future research may also investigate the interaction between physical effort and cognitive effort. For example, research in natural tasks has found that people tend to rely more on memory if the task requires more head movement [47].

It should be noted that the SEEV model served as a qualitative structure (i.e., bandwidth, effort, and saliency) for presenting our results. Our aim was not to compare different sampling models, and it, therefore, remains to be investigated whether the SEEV model yields a better fit to the data than queuing models of visual sampling and other models that use uncertainty and reward/cost of having (in)accurate state information [20], [48]. An inelegance of the SEEV model is that saliency is causally related to bandwidth, because higher bandwidth dials move faster and are, therefore, overall more salient. Furthermore, it is debatable whether bandwidth and value are orthogonal variables, as participants may believe that faster moving (higher bandwidth) dials are also the more important (higher value) dials. We also note that it is difficult to compare the relative contributions (e.g., in terms of variance explained) of bandwidth, effort, and saliency, because bandwidth differs between dials, effort differs between videos, and saliency (e.g., whether a pointer moves fast or not) differs per dial as a function of elapsed time. The criterion with which models can be compared also deserves further examination.

Our results may have various implications for the design of human machine interfaces for the supervisory control of automated processes. In particular, the results make clear how different stimuli conditions and dial configurations compete for attention. For example, we found that it is less likely that an observer gazes to a particular instrument when this instrument requires transition effort, which reinforces the notion that instruments should be within visual reach. It is recommended to investigate whether the present results generalize to more complex tasks, such as supervisory control on-board the cockpit of an aircraft or an automated car, where the operator must not only monitor the instruments but also visually sample competing stimuli of the external environment.

The present work opens up opportunities to provide human operators with real-time feedback when it is predicted that visual and so subsequent task performance may degrade. For example, when a situation is dangerous yet signal indicators do not evidently reveal a danger (e.g., low dial velocity, operator habituated to low signal bandwidth), then that signal could be augmented by temporarily enhancing saliency, to in turn improve task performance.

REFERENCES

[1] P. Hancock, "Automation: How much is too much?" *Ergonomics*, vol. 57, no. 3, pp. 449–454, Dec. 2013.

[2] T. B. Sheridan, *Humans and Automation: Systems Design and Research Issues*. Santa Monica/New York, NY, USA: Hum. Factors Ergonom. Soc./Wiley, 2002.

[3] T. B. Sheridan, "Human centered automation: oxymoron or common sense?" in *Proc. Intell. Syst. 21st Century IEEE Int. Conf. Syst., Man Cybern.*, 1995, pp. 823–828.

[4] M. Donk, "Human monitoring behavior in a multiple-instrument setting: Independent sampling, sequential sampling or arrangement-dependent sampling," *Acta Psychol.*, vol. 86, no. 1, pp. 31–55, Jun. 1994.

[5] K. S. Steelman, J. S. McCarley, and C. D. Wickens, "Theory-based models of attention in visual workspaces," *Int. J. Hum. Comput. Interact.*, vol. 33, no. 1, pp. 35–43, Sept. 2016.

[6] C. A. Gainer and R. W. Obermayer, "Pilot eye fixations while flying selected maneuvers using two instrument panels," *Hum. Factors*, vol. 6, no. 5, pp. 485–500, Oct. 1964.

[7] A. Haslbeck and B. Zhang, "I spy with my little eye: Analysis of airline pilots' gaze patterns in a manual instrument flight scenario," *submitted for publication*.

[8] M. M. Heiligers, T. V. Holten, and M. Mulder, "Predicting pilot task demand load during final approach," *Int. J. Aviation Psychol.*, vol. 19, no. 4, pp. 391–416, Sept. 2009.

[9] A. A. Spady, "Airline pilot scan patterns during simulated ILS approaches," NASA Tech. Paper 1250, Hampton, VA, USA, 1978. [Online]. Available: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19790011511.pdf.

[10] P. M. Fitts, R. E. Jones, and J. L. Milton, "Eye movements of aircraft pilots during instrument-landing approaches," *Aeronaut. Eng. Rev.*, vol. 9, no. 2, pp. 56–66, 1950.

[11] S. J. Landry, "Human-computer interaction in aerospace," in *The Human-Computer Interaction Handbook*, A. Sears and J. A. Jacko, Eds. New York, NY, USA: Taylor & Francis, 2008, pp. 721–740.

[12] J. J. Seeberger and W. W. Wierwille, "Estimating the amount of eye movement data required for panel design and instrument placement," *Hum. Factors*, vol. 18, no. 3, pp. 281–292, Jun. 1976.

[13] J. W. Senders, "The human operator as a monitor and controller of multidegree of freedom systems," *IEEE Trans. Hum. Factors Electron.*, vol. HFE-5, no. 1, pp. 2–5, Sep. 1964.

[14] C. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949.

[15] J. W. Senders, "Visual sampling processes," Ph.D. dissertation, Katholieke Hogeschool Tilburg, Tilburg, The Netherlands, 1983.

[16] N. Moray, "Monitoring behavior and supervisory control," in *Handbook of Perception and Human Performance*, K. Boff, L. Kaufman, and J. Thomas, Eds. New York, NY, USA: Wiley, 1986, pp. 40:1–40:55.

[17] J. W. Senders, J. I. Elkind, M. C. Grignetti, and R. Smallwood, "An investigation of the visual sampling behaviour of human observers," Rep. no. NASA CR-434 , Nat. Aeronaut. Space Admin., Washington, DC, USA, 1966.

[18] L. Fogel, "A note on the sampling theorem," *IEEE Trans. Inf. Theory*, vol. 1, no. 1, pp. 47–48, Mar. 1955.

[19] J. Carbonell, "A queueing model of many-instrument visual sampling," *IEEE Trans. Hum. Factors Electron.*, vol. HFE-7, no. 4, pp. 157–164, Dec. 1966.

[20] J. Carbonell, J. Ward, and J. Senders, "A queueing model of visual sampling: Experimental validation," *IEEE Trans. Man Mach. Syst.*, vol. 9, no. 3, pp. 82–87, Sept. 1968.

[21] T. Sheridan, "On how often the supervisor should sample," *IEEE Trans. Syst. Sci. Cybern.*, vol. 6, no. 2, pp. 140–145, Apr. 1970.

[22] T. O. Kvålseth, "Human information processing in visual sampling," *Ergonomics*, vol. 21, no. 6, pp. 439–454, Jun. 1978.

[23] A. Borji and L. Itti, "State-of-the-art in visual attention modeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 185–207, Jan. 2013.

[24] J. M. Wolfe, "Guided search 4.0: Current progress with a model of visual search," in *Integrated Models of Cognitive Systems*, W. D. Gray, Ed. London, U.K.: Oxford Univ. Press, 2007

[25] J. Najemnik and W. S. Geisler, "Optimal eye movement strategies in visual search," *Nature*, vol. 434, pp. 387–391, Mar. 2005.

[26] D. D. Salvucci and N. A. Taatgen, "Threaded cognition: An integrated theory of concurrent multitasking," *Psychol. Rev.*, vol. 115, no. 1, pp. 101–130, 2008

[27] N. Sprague, D. Ballard, and A. Robinson, "Modeling embodied visual behaviors," *ACM Trans. Appl. Perception*, vol. 4, no. 2, pp. 1–23, Jul. 2007.

[28] C. D. Wickens, J. Goh, J. Helleberg, W. J. Horrey, and D. A. Talleur, "Attentional models of multitask pilot performance using advanced display technology," *Hum. Factors*, vol. 45, no. 3, pp. 360–380, Sept. 2003.

[29] W. J. Horrey, C. D. Wickens, and K. P. Consalus, "Modeling drivers' visual attention allocation while interacting with in-vehicle technologies," *J. Exp. Psychol., Appl.*, vol. 12, no. 2, pp. 67–78, Jun. 2006.

[30] C. Wickens, J. McCarley, and K. Steelman-Allen, "NT-SEEV: A model of attention capture and noticing on the Flight Deck," *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, vol. 53, no. 12, pp. 769–773, Jan. 2009.

[31] M. D. Byrne and R. W. Pew, "A history and primer of human performance modeling," *Rev. Hum. Factors Ergonom.*, vol. 5, no. 1, pp. 225–263, Jan. 2009.

[32] M. D. Fleetwood, "Refining theoretical models of visual sampling in supervisory control tasks: Examining the influence of alarm frequency, effort, value, and salience," Ph.D. dissertation, Rice Univ., Houston, TX, USA, 2005.

[33] Y. B. Eisma, J. C. F. de Winter, C. D. D. Cabrall, and P. A. Hancock, "Situation awareness and overt bodily expression," *submitted for publication*.

[34] J. I. Elkind, "Characteristics of simple manual control systems," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, USA, 1956.

[35] M. Nyström and K. Holmqvist, "An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data," *Behav. Res. Methods*, vol. 42, no. 1, pp. 188–204, Feb. 2010.

[36] K. Rayner, "Eye movements and attention in reading, scene perception, and visual search," *Quart. J. Exp. Psychol.*, vol. 62, no. 8, pp. 1457–1506, Jun. 2009.

[37] SensoMotoric Instruments GmbH, *BeGaze Manual*. Teltow, Germany: SMI, 2014.

[38] *Road Vehicles—Measurement of Driver Visual Behaviour With Respect to Transport Information and Control Systems—Part 1: Definitions and Parameters*, ISO Standard 15007-1, 2014.

[39] H. Godthelp, P. Milgram, and G. J. Blaauw, "The development of a time-related measure to describe driving strategy," *Hum. Factors*, vol. 26, no. 3, pp. 257–268, Jun. 1984.

[40] R. Bootsma and R. R. D. Oudejans, "Visual information about time-to-collision between two objects," *J. Exp. Psychol., Hum. Perception Perform.*, vol. 19, pp. 1041–1052, 1993.

[41] R. D. Morey, "Confidence intervals from normalized data: A correction to Cousineau (2005)," *Tutorial Quant. Methods Psychol.*, vol. 4, no. 2, pp. 61–64, 2008.

[42] C. D. Wickens, "Visual attention control, scanning, and information sampling," in *Applied Attention Theory*, C. D. Wickens and J. S. McCarley, Eds. Boca Raton, FL, USA: CRC Press, 2008, pp. 41–61.

[43] H. Summala, T. Nieminen, and M. Punto, "Maintaining lane position with peripheral vision during in-vehicle tasks," *Hum. Factors*, vol. 38, no. 3, pp. 442–451, Sept. 1996.

[44] A. Popovici and P. M. T. Zaal, "Effects of retinal eccentricity on human manual control," in *Proc. IMAGE Conf.*, 2017.

[45] J. W. Senders, I. B. Webb, and C. A. Baker, "The peripheral viewing of dials," *J. Appl. Psychol.*, vol. 39, no. 6, pp. 433–436, Dec. 1955.

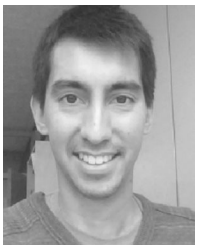
[46] J. W. Senders, A. B. Kristofferson, W. H. Levison, C. W. Dietrich, and J. L. Ward, "The attentional demand of automobile driving," *Highway Res. Rec.*, vol. 195, pp. 15–33, 1967.

[47] D. H. Ballard, M. M. Hayhoe, and J. B. Pelz, "Memory representations in natural tasks," *Memory*, vol. 7, no. 1, pp. 66–80, Dec. 2007.

[48] N. Sprague and D. Ballard, "Eye movements for reward maximization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2003, pp. 1467–1474.

Yke Bauke Eisma received the M.Sc. degree in mechanical engineering from the Delft University of Technology, Delft, The Netherlands, in 2017. Here, he is currently working toward the Ph.D. degree at the Control and Simulation Department.

Here, he is working on the design of an ecological tutoring system for air traffic control operators.



Christopher D. D. Cabrall received the B.A. degree (summa cum laude) in psychology and linguistics with a minor in computer science from Northeastern University, Boston, MA, USA, in 2007, and the M.Sc. degree (summa cum laude) in human factors and ergonomics from San Jose State University, San Jose, CA, USA, in 2010. He is currently working toward the Ph.D. degree at the Delft University of Technology, Delft, The Netherlands.

Within the HFAuto project, he is investigating issues of human automation interaction, driver vigilance, and mode awareness.

Joost C. F. de Winter received the M.Sc. degree in aerospace engineering and the Ph.D. degree in the domain of human factors and ergonomics (cum laude) from the Delft University of Technology, Delft, The Netherlands, in 2004 and 2009, respectively.

His research interests focus on human factors and statistical modeling, including the study of individual differences, driver behavior modeling, multivariate statistics, and research methodology. He is currently an Associate Professor with the Mechanical Engineering Department, Delft University of Technology.