Fixating, attending, and observing: a behavior analytic eye-movement analysis

Steffen Hansen and Erik Arntzen

Department of Behavioral Science, Oslo and Akershus University College, Oslo, Norway

ABSTRACT
The use of eye-tracking technology to study eye-movements has increased substantially over the last decade. For instance, areas that relate to image scanning, matching-to-sample learning, driving, and reading exhibit this trend. Despite the fact that eye-tracking technology reveals a participant’s eye-movement and fixation pattern during an experiment, when can we say that he or she has emitted an ocular observing response to a visual discriminative stimulus? The purpose of the present article is to focus on some influential publications on the observing response and eye-fixation, investigated with eye-tracking technology, and thereby to provide a conceptual distinction between fixating, observing, and attending. Basically, (a) eye-fixations are detected by event-specific algorithms; (b) ocular observing responses occur with and without clear-cut eye-fixation; and (c) ocular observing responses are context-specific, hence, vary across behaviors, settings, and individuals. Finally, we describe in-depth dependent fixation measures as rate, number, proportion, and pattern to offer a broad view on how eye-tracking analysis can provide us with a better understanding of complex human behavior.

With eye-tracking glasses on and carrying a package, Lisa exited the post office. She headed toward our location, took off the eye-tracking glass, and gave us a brief summary of her experience:

Well, as I opened the entrance door to the post office, while looking slightly to the left, a bright, standing object caught my attention. I looked straight at the object and quickly identified it as the ticket number machine. Approaching the machine, I observed two buttons; a green for a ticket to the regular mail line and a red for a ticket to the package pick-up and delivery line. Eager to pick up my package, one of three items that I had previously identified thru a reinforcer assessment, I pressed the red button and a white ticket slid out. Then I sat down and waited for number 117 to pop up on the monitor at the wall in front of me. Although my eyes were wide open and pointing in the direction of the monitor, I found that I was imagining myself unwrapping the package in order to reveal its contents. Suddenly a loud tone caught my attention. I had heard that sound several times before, and almost instinctively I "zoomed in" on the monitor. Number 117 flashed on the screen, while simultaneously instructing the ticket holder to go to service window number 8. My eyes moved slowly from left to right. Window number 2…5…8! Great! Package pick-up and delivery was located at the far right. The clerk greeted me with

CONTACT Steffen Hansen steffenhansen247@gmail.com; Erik Arntzen erik.arntzen@equivalence.net

© 2015 Norwegian Association for Behavior Analysis
a smile, as I handed her my package pick-up slip. Upon receiving the package, I looked for
the exit sign, located it, and made my way back to you.

Lisa then pressed “stop recording.” It was perfect. We were now able to replay the event
thru Lisa’s recorded eye-movements and compare them to her visual experience.

If behavior analysis is to offer a comprehensive understanding of complex human
behavior, as Palmer (2010) noted, it is vital that we increase the resolution of our
experimental procedures. Metaphorically, he suggested that we put the behaviors of
eye-movements under the microscope because they are yet another dependent variable
that has far-reaching implications.

We study eye-movements by using eye-tracking technology. Hitherto, it has been an
infrequently used but invaluable instrument in behavior-analytic research (e.g., Dube
et al., 1999; Kirshner & Sidman, 1972; Tomanari et al., 2007). Data on eye-movements
can augment measures of selection responses by controlling the pattern and duration of
visual discriminative stimuli. Thus, eye-movement data are important in regard to a
wide variety of behavioral phenomena, particularly those at the borders of our cap-
ability to analyze experimentally. For instance, when participants scan complex visual
displays, they are exposing themselves to a sequence of visual discriminative stimuli.

In a matching-to-sample arrangement, to take a second example, longer observation
durations indicate that it is more plausible that any subsequent selection response will
be partially controlled by the sample stimulus and partially controlled by the interven-
ing visual discriminative stimuli. A fixation on a comparison stimulus, without the
evocation of a discriminative response, suggests that this particular comparison stimu-
lus played the role of $S^A$ for its selection. In addition, excessively long response times to
presented comparison stimuli might indicate a cascade of fixations to either one
stimulus or the scanning of several of the comparison stimuli. Therefore, a more
comprehensive understanding of the controlling variables is available if we consider
these inter-trial events.

Furthermore, studies on visualization and imagery might make use of eye-tracking
technology. When asked to visualize a pattern, do people tend to turn away from
distractions and look at a blank wall? Do their eyes move in a pattern that corresponds
to elements of the task? For example, when visualizing the moves of a knight on a
chessboard, do our eyes move across the chessboard in accordance with the moves that
we imagine? For instance, in a behavioral perspective study on listening and auditory
and visual imagining, which was supported by PET and fMRI tasks, Schlinger (2008,
2009) pointed to evidence which indicated that during visual imagery, eye-movements
reflected the perceived movements of the same visual scene (i.e., Laeng & Teodorescu,
2002).

Behavioral utility envisioned, the main purpose of the present article is to discuss
structural and functional components of eye-movements by focusing on publications
that relate to eye-fixation and the ocular observing response, explored with eye-tracking
technology and, thereby, to propose a general conceptual understanding of fixation
events, observing response events, and how these relate to attending. Reviewed articles
and books resulted from keyword searches on observing behavior, eye-fixation, and
eye-tracking methodology on PsychNET and related search engines. While three major
works on eye-tracking methodology (i.e., Duchowski, 2007; Holmqvist et al., 2011;
Horsley, Eliot, Knight, & Reilly, 2014) have led us to other useful studies, in present review, we evaluated articles on eye-fixation based on their relevance to conceptual issues, as well as their applicability in behavioral analytic research.

Mayer (2010) stated that “a serious challenge for eye-tracking researchers is to find the sometimes-missing link between eye-fixation measures and learning outcome (or cognitive performance) measures” (p. 170). By observing relatively stable changes in fixation rates and other eye-movement topographies, as a result of relatively stable changes in environment–behavior relations (i.e., learning), behavioral analysts attempt to address this challenge.

Eye-tracking technology provides us with the opportunity to explore several eye-movement topographies. Yarbus (1967) referred to these topographies as (a) saccades, or rapid eye-movements; (b) smooth pursuits, such as when eyes follow a pendulum movement; and (c) eye-fixations, defined as “sensed visual stimuli that are stationary relative to an observer’s head and eyes” (p. 105). Behavioral analytic researchers have thus been armed with yet another tool to establish control of the variables that govern complex human behavior.

In a typical eye-tracking experiment (e.g., Dube et al., 1999), a participant is equipped with a head-mounted eye-tracking system that consists of two small video cameras, an infrared light, and a double-sided dichroic mirror (see Figure 1). The mirror guides light by selectively transmitting and reflecting different wavelengths, but it appears transparent to the participant. Additionally, the scene camera shows a significant portion of a participant’s field of view, and this is reflected on the outside of the mirror. The eye camera records eye-movements from the reflected image of the eye on the inside of the mirror via a corneal reflection system. Because the image reflection systems are head-mounted, Dube et al. noted that it is not necessary to immobilize a participant’s head. However, other research labs (e.g., Arntzen & Hansen, 2013) have found a non-intrusive chin cup (i.e., a head-support system that voluntarily immobilizes a participant’s head during recording—see Figure 1) useful. Finally, Dube and colleagues analyzed their video signals by using a computer that ran ISCAN Point-of-Regard Data Acquisition software. The output was “a real-time video field-of-view image with a superimposed cursor that indicated the participant’s point of gaze” (Dube et al., 1999, p. 9). The eye-tracking apparatus that was employed by Dube

![Figure 1.](image-url) A participant situated in an experimental preparation. Wearing an eye-tracking glass, recording eye-movements of the left eye, the head is supported by and is rested on a non-intrusive chin cup system.
et al. is also reviewed by Duchowski (2007) and is one of a variety of video-based eye-tracking measures that is on the market today (Duchowski, 2007; Holmqvist et al., 2011; Horsley et al., 2014; Salvucci & Goldberg, 2000; van der Lans, Wedel, & Pieters, 2011).

The use of eye-tracking technology when studying eye-movements in various research settings has increased substantially over the last decade. For example, Salvucci and Goldberg (2000) described this method as a “window into observers’ visual and cognitive processes”. In behavioral analytic terms, it is a method that can measure a person’s ocular observing responses to visual discriminative stimuli. The implementation of eye-tracking technology in behavioral studies on image scanning, matching-to-sample learning, driving, and reading indicates such a trend (e.g., Arntzen & Hansen, 2013; Dube et al., 2006; Duchowski, 2007; Hansen & Arntzen, 2013; Holmqvist et al., 2011; Horsley et al., 2014; Salvucci & Goldberg, 2000; Tomanari et al., 2007). Still, this procedure presents challenges. How do we operationally distinguish saccades from eye-fixations? How do we know that someone has not only looked at, that is, fixated at a visual discriminative stimulus but also observed it? Furthermore, how is attending different from fixating and observing? Finally, is there a point at which we can say that we have perceived an observed event?

The selection of a response depends on how much contact an organism has had with each one of the stimuli involved. Dinsmoor referred to contact as

the impingement of the stimulus energy on the receptor cells of the relevant sensory apparatus, which typically requires or is modulated by auxiliary behavior known as observing (e.g., looking at and focusing on the stimulus object, touching it, tasting it, etc.). (p. 365).

Wyckoff (1952) was the first to define observing behavior. In his dissertation, he wrote that “we shall adopt the term ‘observing response’ (RO) to refer to any response that results in exposure to the pair of discriminative stimuli that are involved” (p. 431). Nonetheless, Donahoe and Palmer’s (2004) gradation that observing responses are “acquired environment-behavior relations whose primary function is to affect the sensing of stimuli, which then function as conditioned reinforcers for those relations” has a greater appeal, as it clearly distinguishes eye-fixations from ocular observing responses. That is, an eye-fixation establishes no more than that an individual is looking straight at a particular stimulus object, whereas an ocular observing response implies subsequent differential responding to that object. Wyckoff (1952) referred to such differential responding as “effective responding,” namely responses “upon which reinforcement is based; that is, running, turning right or left, lever pressing, etc.” (p. 431). An example of attending, fixating, and observing during a potential conditional discrimination task is illustrated in Figure 3. Further along in the process, it brings us to yet another question. Although eye-tracking equipment records potential ocular observing response events, when can we tell that someone has also perceived such events? Behavior analytically, perception is defined as the acquisition of stimulus control, that is, when an organism behaves one way in the presence of a given stimulus and another way in its absence (Baum, 2005). In short, to understand the distinction between attending, observing, and perceiving while fixating, it is helpful to first separate and examine the structural and functional components of fixation events, detected with eye-tracking technology.
Structural and functional components of fixation events

The structural component

To understand the structural component of fixation events, it is necessary to understand the events that we refer to as visual perception and eye-fixation (see Table 1). Building on this foundation, we detect eye-fixations (see Table 2) through the use of preset algorithmic measures (e.g., Salvucci & Goldberg, 2000; van der Lans et al., 2011).

Visual angle and acuity

Visual perception is a complex integration of light, form, contrast, and color sense (Khurana, 2008). Visual acuity refers to the measure of form sense and concerns the thresholds at which we are able to discriminate a visual stimulus spatially. On the other hand, visual angle is a practical way of measuring the distance between two visual reference points, for example, from the center of one visual stimulus to the center of another visual stimulus (Khurana, 2008, p. 39). Structurally, behavioral researchers regard it as an eye-fixation to a visual discriminative stimulus when the point-of-gaze cursor (i.e., the position where the eye looks) is within an area of 2° of visual angle from the center point of a visual discriminative stimulus at a viewing distance of 55 cm. Dube and colleagues (1999, p. 11) used an angle distance of 2°, as it is regarded as the “diameter of the foveal area of greatest acuity” (Bennett & Rabbetts, 1989, p. 18). With this distinction in mind, we shall refer to the sensation of visual discriminative stimuli in which the angle distance of the diameter of the foveal area is 2° or less as clear-cut eye-fixation (Skinner, 1953, p. 124) and, furthermore, the sensation of visual discriminative stimuli in which the angle distance of the diameter of the foveal area is more than 2° as peripheral vision (Duchowski, 2007, p. 11).

| Table 1. Structural events and terms related to fixation event detection. |
|-----------------|------------------|-----------------------------------|
| Structural events | Terms | Description |
| Visual perception | Visual acuity | A complex integration of light, form, contrast, and color sense. The measure of form sense and concerns the thresholds at which we are able to discriminate a visual stimulus spatially. |
| | Visual angle | A practical way of measuring the distance between two visual reference points, for example, from the center of one visual stimulus to the center of another visual stimulus. |
| | Peripheral vision | The sensation of visual discriminative stimuli in which the angle distance of the diameter of the foveal area is more than 2°. |
| | Clear-cut eye-fixation | The sensation of visual discriminative stimuli in which the angle distance of the diameter of the foveal area is 2°. |
| Eye-Fixation | Eye-in-head fixation | Eye is motionless in its socket (i.e., during fixations and smooth pursuits, the head and eyes follow the visual stimulus in a synchronized fashion). |
| | Eye-on-stimulus fixation | Eyes fixate on a visual stimulus but move inside the head (i.e., during fixations and smooth pursuits, the eyes fixate on the visual stimulus, regardless of head movement). |

Note: The table describes the structural events and terms that are involved in fixation events. The column to the left refers to the structural events, the middle column covers terms related to the structural events, whereas the right column describes the events and terms.
Eye-fxation

As an event, the structure of eye-fxation is further delineated by (a) the eye-in-head fxation, which occurs when the eye is motionless in its socket (i.e., during fxations and smooth pursuits, the head and eyes follow the visual stimulus in a synchronized fashion), and (b) the eye-on-stimulus fxation, which occurs when the eyes fxate on a visual stimulus but move inside the head (i.e., during fxations and smooth pursuits, the eyes fxate on the visual stimulus, regardless of head movement) (Holmqvist et al., 2011; Yarbus, 1967).

Event detection algorithms

Event detection algorithms are used in the characterization of eye-movement data sequences. The analysis of such sequences reveals whether or not novel eye-movement patterns have occurred (i.e., saccade or fxation events). Event detection algorithms make use of three specifc sets of information: (a) gaze position (i.e., x, y coordinates on the visual field), (b) gaze velocity (i.e., speed in a certain direction), and (c) gaze acceleration (Holmqvist et al., 2011). Based on this information, investigators have proposed a taxonomy of fxation identifcation algorithms: dispersion-based, velocity-based (Salvucci & Goldberg, 2000), and probability-based (van der Lans et al., 2011).

Dispersion-based algorithms. Dispersion-based algorithms use positional information and can, therefore, be applied to each recorded data point because all of the areas within a person’s visual field are subject to fxation events (Salvucci & Goldberg, 2000). The dispersion-threshold identifcation algorithm (I-DT) identifes eye-fxations through the use of contiguous data samples that are located within a predetermined “window-size.” The literature suggests that an observing response to

<table>
<thead>
<tr>
<th>Table 2. Algorithms related to fxation event detection.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event detection algorithms</td>
</tr>
<tr>
<td>-------------------------------</td>
</tr>
<tr>
<td>Dispersion-based algorithms</td>
</tr>
<tr>
<td>Velocity-based algorithms</td>
</tr>
<tr>
<td>Probability-based algorithms</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Note: The table summarizes the most frequently used algorithms to detect eye-fxation events. The column to the left refers to types of algorithms, the middle column delineates different models of the specific types of algorithms, whereas the right column describes appropriate applicability of the respective algorithms.
a visual stimulus has occurred when contiguous data samples within the predetermined “window-size” equal a duration length that is between 100 and 250 ms (Salvucci & Goldberg, 2000; van der Lans et al., 2011; Yarbus, 1967). For example, a participant wears an eye-tracking glass and scans a picture of a grass field that contains a horse, a cow, a pig, and a hen. Let’s say that we want to know whether or not the participant has fixated on the horse. Hence, we adjust the predetermined “algorithmic window” so that it only includes the horse. Furthermore, the eye-tracker samples eye-movement data points that are separated by 30 ms. Hence, if we operate with a minimum fixation criterion of at least 250 ms, we would need a sequence of at least 9 (i.e., 250 ms/30 ms = 8.3) contiguous data points within our “algorithmic window” to identify a data sequence as a fixation event. In short, the I-DT exploits eye-movement data points of low velocity because they tend to register relatively close to each other.

**Velocity-based algorithms.** Velocity-based algorithms are suitable for high-speed eye-trackers because they gather eye-movement data points at a relatively high sampling frequency. The algorithms recognize fixations as being strings of eye-movement data samples with a maximum velocity that does not exceed the preset threshold (i.e., 10–50 deg/s). The time span is set to no less than 100 ms, based on research that demonstrates that shorter time spans do not allow for an observing response (i.e., Salthouse & Ellis, 1980; Salthouse, Ellis, Diener, & Somberg, 1981). Assuming that the sampling rate is constant, velocities are measured as the distances between the sampled eye-movement data points (Duchowski, 2007; Holmqvist et al., 2011; Salvucci & Goldberg, 2000; van der Lans et al., 2011).

The velocity-threshold identification (I-VT) algorithm is user-friendly because it separates fixation- and saccade-segments based on their point-to-point velocities (Salvucci & Goldberg, 2000; van der Lans et al., 2011). For example, when the speed between two eye-movement data points is lower than 100 deg/s, a fixation is registered, and when the speed is higher than 100 deg/s, a saccade is registered.

**Probability-based algorithms.** The hidden Markov model, I-HMM (Salvucci & Goldberg, 2000; van der Lans et al., 2011), and the binocular-individual threshold (BIT) algorithm (van der Lans et al., 2011) are velocity-based, probabilistic models because they utilize sequential data segments in their computational protocols. Furthermore, the algorithms establish the most probable depiction of an eye-fixation by employing within-variability measures of the velocity distributions for saccade and fixation segments. Based on the analysis of eye-movement data that were collected by different eye-trackers, van der Lans et al. (2011) argued that a probabilistic approach provides a more valid fixation measure than a fixed-threshold approach does. Most fixation thresholds in eye-movement data are fixed a priori across individuals and tasks (van der Lans et al., 2011, p. 240). As a result, the algorithms do not allow for between-subject and within-subject variability. The BIT algorithm, on the other hand, uses velocity thresholds that are based on natural fixation variability for (a) the context, (b) the task, and (c) the individual. The generic nature of a probabilistic approach to fixation identification is similar to Skinner’s (1935) writings on the variable nature of stimulus–response relationships.
Recognizing that an operant is generic in nature (i.e., variability in the stimulus–and response–class relationship) justifies the use of a probabilistic approach to fixation threshold identifications, as this method accounts for an additional number of probable observing response events.

**The functional component**

The functional component of fixation events, in contrast with the structural component, is determined on the causal factors of the occurrence of observing response events as defined earlier. Hence, a fixation event is considered an observing response, when it occurs because it has led to effective responding (see Figure 3). The main advantage of a functional approach to eye-movements is that it provides us with opportunities to obtain behavioral dependent fixation measures, or observing responses, as number, rate, pattern, and proportion. Before turning our eyes toward such measures, figuratively speaking, we will first review publications that experimentally distinguish among the concepts of attending, looking, observing, and perceiving (see Figure 2).

**Figure 2.** Based on reviewed literature, a flowchart illustrates the controlling relations and behavioral principles that govern an eye’s contributing measures to complex human behavior with attending and looking at one end of a continuum and observing and perceiving further along the continuum, respectively.
Attending and looking

Skinner (1953) wrote that:

Just as we may attend to an object without looking at it, so we may look at an object without attending to it. We need not conclude that we must then be looking with an inferior sort of behavior in which the eyes are not correctly used. The criterion is whether the stimulus is exerting any effect upon our behavior. When we stare at someone without noticing him, listen to a speech without attending to what is said, or read a page "absent-mindedly," we are simply failing to engage in some of the behavior which is normally under the control of such stimuli. (p. 124)

Dinsmoor (1985) was skeptical of the idea that attending could be seen as a separate concept from that of observing. As solid evidence forced him to accept the idea, he speculated that attending had to do with "analogous processes occurring further along in the sequence of events, presumably in the neural tissue" (p. 365). His proposal was grounded in a distinguished experiment by Jenkins and Harrison (1960). They exposed pigeons to a tone of 1000 Hz on a continuous basis during the conditioning of pecking. Throughout subsequent test periods, the birds showed no variations in response rates in the presence of the tone of 1000 Hz or of tones of higher or lower frequency. Consequently, in another group of birds they reinforced key pecking in the presence of a tone of 1000 Hz (S\textsuperscript{R}) but not in its absence (S\textsuperscript{A}). Results indicated a steep, symmetric generalization gradient around an apex of 1000 Hz, suggesting that the
tone had acquired substantial evocative power. As a result, Dinsmoor acknowledged that “the existence of some process of a more central nature, which might appropriately be called attention” (1985, p. 371) and, furthermore, that this process might develop in accordance with the same behavioral principles that describe the acquisition and maintenance of observing responses. Rudolph and Houten (1977) agreed, stating that “the tone may inform the subject that the environment containing the possibility of reinforcement is present” (p. 330).

From auditory to ocular attending, Skinner’s (1953) writings were consistent with Jenkins and Harrison’s (1960) finding:

If the light begins to flicker while the pigeon is looking elsewhere, the flicker is seen at one side of the visual field. The behavior of looking directly toward the light is then optimally reinforced. We say that the light “captures the undivided attention” of the bird. (p. 123)

Skinner (1953) further argued that attending is not a form of behavior; he claimed that it is “a controlling relation—the relation between a response and a discriminative stimulus” (p. 123). Skinner also noted that someone who pays attention is under special control of a stimulus and that this relationship is easily detected when receptors are directly oriented toward the stimulus; however, as he emphasized, this orientation is not a necessity: “an organism is attending to a detail of a stimulus, whether or not its receptors are oriented to produce the most clear-cut reception” (p. 124). Data obtained by Arntzen and Hansen (2013) support this view as the researchers showed that participants often attended to and mouse-clicked a comparison stimulus before their eyes had moved and fixated directly on that specific comparison stimulus.

Observing and perceiving
Salthouse and Ellis (1980) reviewed studies on measures that accompany the functional component of an entire eye-fixation. For example, in a psycholinguistic approach, Smith (1971) speculated that perception of a visual stimulus required an observing response lasting approximately 250 ms. Almost 100 years earlier, Dodge (1907) experimentally demonstrated that stimulus discrimination, or perception, required a minimum of 100 ms. By using an “escapement exposure apparatus, in which each new exposure produced by the rapid movement of the words into place behind a narrow slit” (p. 46), Dodge presented words for 48 ms, 70 ms, 100 ms, as well as for longer periods. Pre- and post-disrupting second stimuli were mirror images of the word “explanation.” Participants did not perceive all of the words that had exposure times of 48 and 70 ms, but they recognized all of the words that had exposure times of 100 ms and above. Referring to inconsistencies in the literature, Salthouse and Ellis (1980) argued that there was little agreement about the minimum fixation time required to produce an observing response and, hence, to perceive a visual discriminative stimulus. Consequently, the authors decided to separate and allocate the contributing measures of the structural and functional components of an eye-fixation.

A thorough investigation was initiated in which four variables were explored. The variables investigated were (a) observing response duration, a functional component; (b) the relative emphasis on speed or accuracy, a structural component; (c) the sequential dependencies across successive observing responses, a functional component;
and (d) the amplitude of the preceding and following saccades, a structural component. Four participants explored a sequentially arranged stimulus arrangement of five letters, which were located in the same spatial location for every trial. Fifty percent of these stimulus arrays contained a single vowel that was randomly placed among the consonants. The minimum time that was necessary to complete an observing response to a simple visual stimulus (i.e., a vowel) was defined to be the length of time at which correct identification (i.e., perception) would occur for approximately 95% of the trials that were presented in a single block. Participants were able to identify a vowel, when the sequence was present for approximately 100 ms. The authors suggested, therefore, that 100 ms was sufficient for participants to distinguish a simple visual stimulus, a vowel, from other simple visual stimuli (i.e., consonants). Interestingly, reviewed literature of the structural component of an observing response suggested minimum duration estimates of 250 ms. Thus, between the structural and functional components, Salthouse and Ellis (1980) observed a discrepancy of 150 ms.

Perplexed by the discrepancy between the structural and functional allocations of the fixation duration, Salthouse et al. (1981) decided to implement a systematic replication of the study by Salthouse and Ellis (1980). Specifically, they tested (a) whether the functional component’s minimal share (i.e., 100 ms) of an entire observing response (i.e., 250 ms) was a result of saccadic suppression (i.e., suppression of an observing response prior to and following an eye-movement) or (b) whether previous estimates of observing response durations were underestimated because of researchers’ inability to develop equipment that could identify more complex levels of perception. Salthouse et al. (1981) believed that these levels were higher-order and required extended observational responding (p. 612).

Hence, Salthouse et al. (1981) created three experiments that would either replicate or fail to replicate the previous findings. In the first experiment, two participants tested the relative effectiveness of observing responses to alphabetic characters during all of the segments of the fixation period (p. 612). The researchers confidently rejected the saccadic suppression interpretation because all of the segments of the entire eye-fixation indicated observational responding to the alphabetic characters. The second and third experiments confirmed that extended observational responding was occurring. Specifically, Salthouse et al. (1981) found that the observing response time increased after correct effective responding reached an asymptote. Hence, they speculated that these two observing response measures (i.e., before and after the asymptotic level) could be suitable as dependent variables in the investigation of extended observing. The investigation was accomplished by presenting a second alphabetic character while the first character was observed. The investigators reasoned that this would temporarily interrupt the observing response and cause a lengthening of the observing response time. This period of extended responding could possibly offer an estimate of the time course of prolonged observing responses. Indeed, changes in the observed character did increase the duration of an observing response. Salthouse et al. (1981) concluded, therefore, that the discrepancy between the total time of an observing response and the entire duration of an eye-fixation was in fact minimal.
**Dependent measures**

**Number**

Fixation number is measured in three different ways, referred to as (a) fixation density (Henderson, Weeks Jr., & Hollingworth, 1999), (b) fixation frequency, and (c) fixation latency (e.g., Duchowski, 2007; Holmqvist et al., 2011; Horsley et al., 2014). Fixation density is typically measured when researchers want to count the number of fixations in a narrowed area of interest in the visual field, regardless of fixation durations. Fixation frequency is a count of the entire number of fixations within an individual’s visual field. Fixation latency is measured in two ways: (1) as the total number of observing responses to visual stimuli between stimulus onset and task completion (e.g., the number of observing responses to words on one page), or (2) in a matching-to-sample preparation (see Figure 3), as the number of observing responses to visual sample and comparison stimuli per selection response (considered to be a measure of the strength of stimulus control).

Fixation number has proved to be a reliable dependent measure in matching-to-sample arrangements. For instance, Dube et al. (2006) studied observing responses as a function of two levels of complexity: two or four sample stimuli that were presented simultaneously in a multiple sample, delayed matching-to-sample arrangement. It was shown that an increase in the number of simultaneously displayed sample stimuli did not influence the average number of fixations to each presented sample stimulus. Likewise, Tomanari et al. (2007) studied observing responses in a two-stimulus discrimination arrangement with both eye-movements and manual responses (i.e., mouse-clicking for S_D or S_A stimuli) as the observed responses. Results showed that participants looked at visual S_D and S_A stimuli at a higher rate than they mouse-clicked these same visual stimuli.

**Rate**

Fixation rate is defined as the number of fixations, or observing responses (see Figure 3), during a certain time period or a certain task completion, or the number of fixations, or observing responses, per trial (e.g., Duchowski, 2007; Holmqvist et al., 2011). A high rate of fixations/observing responses per trial is typically seen in the initial training blocks during conditional discrimination training; this rate decreases as certain sample stimuli acquire stimulus control over the selection responses (e.g., Arntzen & Hansen, 2013). This is in accordance with Dinsmoor (1985), who noted that “we see that the proportion of time spent observing the stimulus increases under the same conditions as those producing an increase in control” (p. 369). Thus, we register an increase in fixation events/durations to an accurate selection response and, simultaneously, a decrease in fixation events/durations to inaccurate selection responses.

In a study by Nakayama, Takahashi, and Shimizu (2002), participants solved math problems and spoke aloud during their calculations. Correct observing responses were negatively correlated with task difficulty. Hence, a high number of correct observing responses in a given time period indicated that the mathematical tasks were easy (i.e., tight stimulus control) and vice versa. Therefore, before judging the results of a study, it is important to note whether the study uses the rate of terminal observing responses (i.e., observing responses that result in effective responses that are also regarded as
terminal selection responses, as opposed to effective responses that lead to continued search—see Figure 3) during a certain time period or whether it uses the rate of all observing responses (i.e., within-trial fixation events in addition to the final fixation event that results in a terminal selection response) for a given trial. The strength between the observing response and the terminal effective response is tight when the rate of correct observing responses is high during a certain time period and, moreover, when the rate of observing responses in a given trial is low—it is an indication of tight stimulus control between a certain sample stimulus and its correct selection response.

Pattern

Stimulus control is indicated by decreased variability in the fixation pattern. Dube et al. (2006) examined fixation patterns during matching-to-sample performance with four sample stimuli. Interestingly, as accuracy scores improved, fixation patterns changed from a random fixation pattern to a clockwise fixation pattern (i.e., a decrease in variability). They concluded that additional research would be necessary to identify the variables that control these pattern changes. Additionally, Vakil, Lifshitz, Tzuriel, Weiss, and Arzuouan (2011) asked individuals with and without intellectual disabilities to solve conceptual and perceptual analogies. A conceptual analogy consisted of four pictures. For instance, the top row included a picture of a train on the left side and a railway on the right side, and the bottom row showed a picture of a bus on the left side and a missing picture of a road on the right side. Participants were to choose the correct picture of a road among the four alternatives. Perceptual analogies were presented in the same manner (e.g., perceiving what type of cup is missing among different types of cups). The results for both groups indicated a higher number of within-trial observing responses (i.e., observing responses that lead to additional search and not a terminal selection response) while solving perceptual analogies. Additionally, intellectually disabled individuals made more switches (i.e., within-trial observing responses) while solving perceptual analogies than typically developing individuals did; however, they were less accurate (i.e., more variability in fixation pattern).

Horsley et al. (2014) provided examples on how instruction could influence observing response pattern. First, Buswell (1935) showed that fixation patterns differed between viewing a photograph of the Tribune Tower in Chicago, first without instructions and then with prior instructions given—for instance, look for a face in one of the windows (p. 21). Second, Yarbus (1967) had an individual view the phrase “Repin’, They did not expect him” seven times, each time with different instructions (p. 21). As a result of differences in instructions, observing responses varied notably. The points fixated matched those that provided information with relevance to the instructions given.

West, Haake, Rozanski, and Karn (2006) noted that pattern analysis, also referred to as sequence or scanpath analysis, was not as common as other eye movement measures because the correct tools for this analysis were not integrated into the most common eye movement software (p. 149). Hence, the same authors promote “eyePatterns,” as it is a tool that identifies similarities in fixation patterns, as well as between the experimental variables that can influence their characteristics.
**Proportion**

When investigating the “proportion of eye-fixations,” one compares the number of fixations between areas of interest. Adolphs et al. (2005) worked with a patient with amygdala damage. She showed impairment in her ability to perceive fear from facial expressions. By using eye-tracking technology, Adolphs et al. were able to demonstrate that her deficiency was rooted in an absence of fixation events to the eye region of facial expressions—the region that was regarded as the most important feature of fear recognition. Compared to other areas of the face, the patient rarely fixated at the eye region. Consequently, Adolphs et al. explicitly instructed the patient to look at the eyes. With an increased proportion of eye-fixations allocated to the eye region, according to the authors, the patient’s perception of fearful faces returned to normal.

**Observing response duration: a context-related measure**

**Reading, scene viewing, and visual search**

Rayner (1998, 2009) reported statistics on observing response duration for reading, visual search, and scene viewing. Mean ranges of observing responses were 225–250 ms for reading, 180–275 ms for visual search, and 260–330 ms for scene viewing. Similar observations were reported by van der Lans et al. (2011) who noted significant variability in observing response durations between individuals, stimuli, and tasks. This variability, they argued, was a result of variation in algorithms and fixation threshold settings.

Durations of observing responses have been found to be related to familiarities in and the complexities of the environment. For example, words that seldom appeared in a text were subject to longer fixation durations because they required longer time to produce an observing response (Rayner, 1998). Furthermore, fixation durations were longer in participants who were presented with more complex reading material, which suggested that these visual stimuli required extended ocular observing before producing a correct response (Rayner & Pollatsek, 1989).

The apparent variability in findings for reading suggests the need for sub-dependent measures that address various components of reading material. This approach may also aid behavior analytic interventions for reading. Hence, in addition to reading acting as a unit of analysis, we could compare observing response durations when stimuli are composed of (a) nonsense words, (b) foreign phrases, (c) evocative words, (d) familiar words, (e) proper nouns, etc. Similarly, observing response durations for visual search vary considerably with regard to the complexity of the task (e.g., finding a needle in a hay stack is extremely difficult when compared to finding an elephant in a swimming pool). Thus, sub-dependent units of analysis that are related to the difficulty of the material will greatly improve our understanding of these measures. In regard to scene viewing, sub-dependent units of analysis can include fixation events at a traffic intersection during (a) morning rush, (b) noon traffic, and (c) afternoon rush. In short, splitting the three broad visual discrimination conditions into sub-dependent functional units will multiply the amount of valuable, concrete information that observing responses can provide.
Specific cases

Specialized skills
Expertise (or tight stimulus control) is correlated with longer fixation durations. For instance, Nodine, Locher, and Krupinski (1993) showed that individuals who were professionally involved in chess, darts, and goal keeping engaged in longer observing responses during a game than individuals who were not professionally involved. Nodine et al. speculated that with the improvement of a certain skill, a person would also be able to extract more information from a single eye-fixation per observing response (i.e., make an eye-movement more economic). Behavioral analysts would argue for a more parsimonious explanation. Hence, individuals with a professional background probably engage in extended observing responses because this behavior has a history of reinforcing consequences.

Schizophrenia and Alzheimer’s disease
Research on individuals with schizophrenia and Alzheimer’s disease has suggested that neurological impairments correlate with longer fixation durations (e.g., Lueck, Mendez, & Perryman, 2000). One wonders, however, whether such individuals were attending to the task at hand or just looking straight ahead (see Figure 2). Ishizuka, Kashiwakura, and Oiji (1998) measured fixation durations and delusional talk, and the results indicated a significant correlation between the severity of disturbed speech and an increase in fixation duration. Again, long fixation durations do not always indicate that an individual engages in an extended observing response and, thus, perceives an event. Rather, when looking without engaging in observing responses, the participant is most likely engaging in a competing behavior (i.e., disturbed speech) as a result of reinforcing consequences.

Intellectual disabilities
Dube and colleagues used eye-tracking equipment to examine the relationship between observing behavior and stimulus over-selectivity in intellectually delayed individuals (Dube et al., 1999, 2003). They concluded that stimulus over-selectivity consisted of failures to observe all of the relevant stimuli, as well as short fixation durations to the sample stimuli, that is, insufficient time to engage in observing responses. Hence, Dube et al. (2010) decided to perform a systematic replication in which their goal was to change the experimental procedures to cause a decrease in observing failures, an increase in fixation durations and, as a result, higher accuracy scores as well as the elimination of restricted stimulus control. Four normally capable individuals and 10 individuals with intellectual disabilities (ID) participated in the two-sample delayed matching-to-sample arrangement. Independent measures included the prompting and differential reinforcement of eye-fixations to all the sample stimuli and minimum required eye-fixation durations. The dependent measure was the number of correct responses to comparison stimuli. Eye-movement data indicated that such interventions eliminated observing failures and engaged in longer eye-fixation durations. As a result of an increase in the strength of stimulus control of observing responses to both sample stimuli, as well as longer fixation durations, 8
of 14 individuals obtained high accuracy scores and the remaining 6 participants achieved intermediate accuracy scores.

Finally, a study by Vakil and Lifshitz (2012) revealed different eye-movement patterns in adults with and without Down’s syndrome. Participants solved the Raven Progressive Matrices and typically developed adults engaged in shorter observing response durations for the visual puzzles than did adults with Down’s syndrome. However, both groups engaged in longer observing response durations on occasions when they answered correctly.

The present analysis of the literature reveals that observing response durations are context-related, as they vary substantially across individuals, tasks, and settings. Thus, in line with the conclusions of Dube et al. (2010), eye-tracking experiments on ocular observing response durations suit single-case research designs.

In single-case research designs, the potential utility of eye-tracking equipment is vast. For instance, by connecting a remote monitor (e.g., by using TeamViewer) to a real-time video field-of-view image, experimenters can follow a participant’s eye-movements while he or she is reading, or while matching comparison stimuli to sample stimuli. It furthermore allows for opportunities to deliver immediate positive consequences for successive approximations to effective reading patterns and to recalibrate eye-tracking equipment when necessary.

**Conclusion**

Tracking eye movements is an important additional measure in the study of complex human behavior. We have explored the subcomponents of ocular observing events and, thus, obtained a general understanding of eye-movements and ocular observing response topography. In addition, we are closer to answering our initial question: When has a participant observed an event long enough to produce an effective response that results in reinforcement (i.e., perceived an event)? Literature suggests that observing response duration is context- and task-specific, as well as highly individual. We suggest, therefore, that an individual has observed a visual discriminative stimulus long enough to engage in a reinforcer-producing response (i.e., perceive that stimulus) when the visual discriminative stimulus reliably causes that individual to respond in accordance with the experimenter-defined contingencies.

Furthermore, we point to evidence which suggests that attending, looking, observing, and perceiving operate on something of a functional continuum, with attending and looking—or vice versa—at one end of the scale, with differentially reinforced ocular observing responses further along, and with perceiving at the other end of the continuum (see Figure 2). Thus, attending constitutes a controlling relationship between the visual contact that meets the eye and a visual discriminative stimulus, established and maintained by conditioned reinforcement (see Figure 3). An ocular observing response is an eye-sensation that results in visual access to the discriminative stimuli involved and, as a result, causes an individual to engage in an effective, that is to say differentially reinforced, response.

Eye-tracking technology expands our understanding of visual discrimination by providing us with fixation measures that allow us to study behavioral phenomena that are at the borders of our capability to experimentally analyze. In addition, with
the production of improvements to eye-tracking analysis software, such as “eyePatterns”, it is possible to explore more complex and temporally extended aspects of eye movements, such as fixation patterns.

Dependent fixation measures, including number, rate, and pattern, are relevant to behavioral analytic research because they are discrete events that we can identify in complex visual displays. Dispersion-based algorithms are appropriate for this type of research because eye-movement data segments in complex visual displays, due to their low velocity, tend to cluster relatively close to each other. In order to identify eye-fixation events with dispersion-based algorithms, it is necessary to establish experimenter-defined fixation duration thresholds, which is arguably a limitation to the method. However, this is not an obstacle as long as the selected duration threshold is held constant during all of the phases of an experiment or project.

If behavioral analysts find fixation measures and ocular observing response topography useful in their own line of research, it is recommended that they contact behavioral analytic researchers with related experience. At present, a sampled review of literature on eye-fixation and ocular observing behavior shows that eye-tracking technology makes it possible to study the dependent variables that are embedded in eye-movements. Furthermore, a better conceptual understanding of an eye-fixation, in relation to attending, observing, and perceiving, should extend and enrich our behavioral attempts to explain complex human behavior.

**Acknowledgment**

We thank Robert Mellon, PhD, for his valuable suggestions and comments.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**References**


