Increased Knowledge about Eye Movements. A Systematical Manipulation of Training Directionality in Matching-to-Sample Tasks

Steffen Hansen

Dissertation for the degree of philosophiae doctor (PhD)
Department of Behavioral Science
Faculty of Health Sciences
Oslo and Akershus University College of Applied Sciences

Spring 2017

Avhandling 2017 nr 5

ISSN 1893-0476
ISBN 978-82-8364-050-2

HiOA,
Læringssenter og bibliotek,
Skriftserien
St. Olavs plass 4,
0130 Oslo,
Telefon (47) 64 84 90 00

Postadresse:
Postboks 4, St. Olavs plass
0130 Oslo

Adresse hjemmeside: **http://www.hioa.no/Om-HiOA/Nettbokhandel**
For elektronisk bestilling klikk Bestille bøker
Opplag trykkes etter behov, aldri utsolgt

Trykket hos Allkopi
Trykket på Multilaser 80 g hvit

Acknowledgements

I am honored to express my sincere appreciation to Professor Erik Arntzen for supervising me through this challenging project. In times where steady and rock-solid guidance were needed the most, he stayed course and lead the way. For that I am very thankful!

I am in debt to Kim Liland, Aleksander Vie, and Pedram Sadeghi, for their knowledgeable assistance during the initial phases of our eye-tracking experiments and eye-movement analyses. I am thankful to my research colleagues and fellows, Torunn Lian, Christoffer Eilifsen, Hanna Steinunn, Richard Nartey, Anette Brogaard, Vanessa Ayres, Justice Mensah, Guro Dunvoll, and good friend and flat mate Felix Høgnasson, for their encouragements along the way—the same goes for master and bachelor members of the Complex Human Behavior Lab and friends and colleagues in the Department of Behavioral Science.

I would like to thank my “Norwegian family”—my twin brother Jess Hansen and his fiancée Anette Tosterud, my nephews Sander and Tobias, and “our” dog Salto—for always being there for me.

I am also thankful to my families in Denmark and USA—for always showing their support from abroad.

The list is endless; I am thankful to so many others that have crossed my path during the last four years…

Abstract

The purpose of the first study was to operationalize eye-movement behavior during conditional discrimination training and testing for equivalence class formation and, furthermore, to provide a conceptual systematic framework on visual perception from a behavior analytic viewpoint. Based on influential publications on the observing response and on eye-fixations, we offer a conceptual distinction between fixating, attending, and observing—towards visual perception. Basically, (a) ocular observing responses occur with and without clear-cut eye-fixation; and (b) ocular observing responses are context-specific, hence, vary across behaviors, settings, and individuals. In behavioral research, fixation measures such as time, rate, number, and pattern have profound implications as they reveal important information about eye-movement behavior during the response delay. In study 2, we explored the differential effects of training structures on fixation time and fixation rate during the formation of six 3-member equivalence classes—prepared in a serialized training arrangement. Within-subject designed, nine university-college students participated in the study. Results showed that one of three participants, prepared with MTO, OTM and LS, respectively, responded in accordance with stimulus equivalence. Further, participants who formed equivalence classes revealed longer fixations to sample stimuli and shorter fixation durations to comparison stimuli. Participants fixated both longer and more often at correct comparison stimuli, regardless of equivalence class formation. In study 3, the purpose was to systematically replicate study 2, that is, to explore the differential outcomes in fixation time and fixation rate during the formation of five 3-member stimulus equivalence classes—this time introduced in a group design and a more solid, concurrent training format. Thirty university-college students participated and results replicated findings from study 2: Participants who formed equivalence classes revealed in general longer fixation times to sample stimuli during training and longer fixation times and fixation rates to correct comparison stimuli, regardless of demonstrating equivalence class formation. In addition, fixation rate during training and testing was noteworthy higher for participants prepared with the MTO structure.

Key Words: Attending behavior, complex human behavior, conditioning, eye-fixation, eye-movements, eye-tracking, fixation measures, observing response, stimulus control

Sammendrag

Formålet med den første studien var: (a) å operasjonalisere øyebevegelsesatferd under betinget diskriminasjonstrening og test for respondering i henhold til stimulusekvivalens; og (b) å gi et konseptuelt systematisk rammeverk på visuell persepsjon fra et atferdsanalytisk ståsted. Basert på innflytelsesrike publikasjoner om oppmerksomhet og øyefiksering, fremlegger vi et konseptuelt skille mellom øye-fiksering, oppmerksomhet og observering—som bakgrunn for forståelse av visuell persepsjon. I utgangspunktet, (a) observeringsresponser forekommer med og uten entydig øye-fiksering; og (b) observeringsresponser er kontekstspesifikk; dermed varierer de mellom atferd, miljø og individ. Innen atferdsforskning har øyefikseringsmål som tid, antall og mønster stor betydning, ettersom de avslører viktig informasjon om øye-bevegelsesatferd i tidsrommet mellom trykking på utvalgsstimulus og valg av sammenlikningsstimulus. I den andre studien utforsket vi ulike effekter av treningsstrukturer på øyefikseringstid og antall av øyefikseringer under dannelsen av seks 3-medlems ekvivalensklasser—forberedt i et sekvensielt treningsoppsett. I et inne-deltaker design deltok ni universitetsstudenter. Resultatene viste at en av tre deltakere, forberedt med henholdsvis MTO, OTM og LS responderte i henhold til stimulus ekvivalens. Deltakere som dannet ekvivalensklasser viste til lengre øyefikseringer til utvalgsstimuli og kortere øyefikseringstider til sammenligningsstimuli. Deltakerne fikserte både lengre og oftere på korrekte sammenligningsstimuli, uavhengig av dannelse av ekvivalensklasser. I den tredje studien var hensikten å gjøre en systematisk replikasjon av studie 2, det vil si, og utforske forskjellige utfall i øyefikseringstider og i antallet av øyefiksering under dannelsen av fem 3-medlems stimulusekvivalensklasser—denne gangen introdusert i en gruppe design med et mikset treningsopplegg. Tretti universitetsstudenter deltok i studien, og resultatene replikerte funnene fra studie 2: Deltakere som dannet ekvivalensklasser viste generelt (a) lengre øyefikseringstider til utvalgsstimuli under trening, (b) lengre øyefikserinstider til korrekte sammenligningsstimuli, og (c) et høyere antall øyefikseringer til korrekte sammenligningsstimuli, uansett dannelse av ekvivalensklasser. I tillegg var antall av øyefikseringer under både trening og test bemerkelsesverdig høyere for deltakerne forberedt med MTO-strukturen.

Nøkkelord: Betinget diskriminasjonstrening, eye-tracking, kompleks menneskelig atferd, oppmerksomhetsatferd, observeringsrespons, stimuluskontroll, øyebevegelser, øyefiksering, øyefikseringsmål

Table of Contents

List of Tables and Figures

Introduction

Complex Human Behavior

Complex human behavior, as conceptual, categorical, and symbolic behaviors, has intrigued psychologists and behavioral researchers for at least a century (e.g., Hull, 1920; Keller & Schoenfeld, 1950; Lakoff, 1987; Laurence & Margolis, 1999; Palmer, 2002; Rosch, 1999; Watson, 1930; Zentall, Galizio, & Critchfield, 2002).

Conceptual, categorical, and symbolic behavior. For instance, Keller and Schoenfeld have contributed immensely to our behavior analytic understanding of conceptual behavior. Moreover, they avoid the fallacy of creating cognitive and hypothetical constructs, noting that, "one does not have a concept [...] rather, one demonstrates conceptual behavior, by acting in a certain way" (1950, p. 186). They further explain that such behaviors occur when we respond the same way to a set of different stimuli—the stimuli become members of a class to which we react identically. As an example of *primary stimulus generalization* (e.g., Green & Saunders, 1998), participating in a quiz, I am shown twenty pictures with faces of boys and men from different continents and with different racial backgrounds, and I am asked whether the pictures show males or females? Demonstrating conceptual behavior, my identical response to all the different boys and men is, of course, that they are males—appropriately referred to as a *feature class*, as all demonstrated features of “maleness” (e.g., Fields & Reeve, 2000; Green & Saunders, 1998). Although Hull (1920) and Smoke (1932) initiated experimental explorations on concept formation, Keller and Schoenfeld are credited with the following widely accepted definition: The essence of concepts is "generalization within classes and discrimination between classes" (1950, p. 186). In further realization, Keller and Schoenfeld extrapolated on this definition: “now we can see that equivalent stimuli is what we mean when we speak of a concept” (1950, p. 189).

Paired associates versus symbolic learning. Linguistic symbols are categorized into restricted paired associates, or rote-learned associations, and rich linguistic symbols, or true words (Wilkinson & McIlvane, 2001). Paired associates are typically seen in the context where learning takes place and they characterize non-generative, rote-learned associations among particular stimuli. On the other hand, symbolic relations demonstrate flexible characteristics, as demonstrated in children that learn words at a phenomenal speed and extend these words to different referents and, furthermore, relate these words to other words (e.g., Deacon, 1997; Skinner, 1957)—they illustrate an “understanding of the abstract nature of the relationship between a word and its related class(es) of items/events”(Wilkinson & McIlvane, 2001, p. 356).

Thus, Symbolic behavior is best defined as the understanding of a relationship between a sign and its referent, meaning that the sign is substitutable for its referent in changing contexts; yet, they are not the same thing ((Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979). Sidman concluded that:

> “this treatment of linguistic forms as equivalent to their referents permits us to listen and read with comprehension, to work out problems in their absence, to instruct others by means of speech or text, to plan ahead, to store information for use in the future, and to think abstractly—all of these by means of words that are spoken, written, or thought in the absence of the things and events they refer to” (1994, p. 3).

Examples of symbolic behavior. From the minute that we wake up, we are surrounded by symbols and we often experience a physical reaction to these symbols. For instance, if I get a message from a loved one and the message includes a heart-kissing smiley, then I would most likely feel euphoria—I would feel really good. On the other hand, if I get a message from a loved one and the message contains a heart that is separated into two parts, that is, a broken heart, then I would most likely feel pain—I would feel really bad. This usage of symbols that they actually become the entities which they are supposed to symbolize is profoundly ingrained in our nature (e.g., Sidman, 1994); we react to the sign-symbol of a heart-kiss as if our loved one kissed us for real—by feeling euphoria—and we react to a sign of a broken heart as if our loved one left us for good—by feeling pain.

In another, rather famous, example in which people reacted to a non-language symbol as it was the thing that it characterized was televised worldwide on the news. Responding to the destruction of some American flags, mob members met up in order to plan counterattacks. The flag burners knew that the symbolic act would be interpreted as war, although the act of burning a flag did not harm anyone in any way; no individual was physically touched and no constructions were damaged. Still, the symbolic effect of the event caused a reaction similar to a reaction based on real events (Sidman, 1994).

Eye-Movements

A distinguished scientist proposed that "an examination of eye movements […] might help us to formulate more complete accounts of complex human behavior" (Palmer, 2010, p. 37)—an intriguing proposal, as the head of the complex human behavior lab, which I attended, recently had acquired eye-tracking equipment to study eye-movements in experimental setups pertaining to behavioral complexity.

The experimental study on eye-movements is not a new field (e.g., Kirshner & Sidman, 1972; Schroeder & Holland, 1968, 1969). For instance, Schroeder and Holland concluded “that an eye movement can act as an operant controlled by its consequences. Operant control of eye movements has important implications for human factor analysts concerned with "attention"” (1968, p. 161).

Eye-movement topography. In addition, Yarbus examined eye-movements extensively and operationalized them into three topographical categories, namely, (a) saccades, or rapid eye-movements; (b) smooth pursuits, as when eyes trail a pendulum movement; and (c) eye-fixations, the “sensed visual stimuli that are stationary relative to an observer’s head and eyes” (1967, p. 105). Such topographical distinctions provide researchers with the opportunity conceptualize new dependent measures, for instance, (1) the duration of eye-movements in a certain direction, (2) the duration of an eye-fixation within a certain area, (3) the number of eye-fixations within a certain area, and (4) the pattern and sequence of eye-movements across settings and as a function of experience (e.g., Duchowski, 2007; Hansen & Arntzen, 2015; Horsley, Eliot, Knight, & Reilly, 2014).

Eye-tracking. Hence**,** research on eye-movements and fixation measures have occupied scientists for more than a century (e.g., Dodge, 1907). Nevertheless, with a burst in manufacturing of eye-tracking equipment, the interest in eye-movements has experienced a renaissance. Now, a variety of disciplines take advantage of eye-tracking technology—among others fields, marketing research, medical research, representatives for advancement in sports performance, and education (e.g., Duchowski, 2007; Holmqvist et al., 2011; Horsley et al., 2014). In behavioral research, this innovative method is taking off as well, as it demonstrates its utility in studies on attending behavior.

Attending behavior. If we are to learn any planned behavior, for example, to name the picture-letters in the alphabet or, even better, to read words in different languages (e.g., “moose” in English or “elg” in Norwegian) with comprehension, then it is crucial that we, first of all, look at the stimuli involved. As Skinner wrote:

> “Looking and listening are forms of behavior, and they are strengthened by reinforcement. A pigeon can learn to match colors, for example, only if it "pays attention to them." The experimenter makes sure that it does so, not by attracting its attention, but by reinforcing it for looking” (Skinner, 1961, p. 182).

When we look at, or attend to, a stimulus that is presented in a learning-contingency, it is more likely that we learn the behavior involved. Eye-tracking technology and operational definitions of eye-movement topographies, as *saccades* and *eye-fixations,* provide behavioral researchers with the tools to establish what an individual is looking at and, thus, arrange for consequences for such behavior. Attending behavior, with the means of eye-tracking equipment, has been explored in areas as image scanning, reading practice, and conditional discrimination procedures (e.g., Arntzen & Hansen, 2103, April; Dube et al., 2006; Dube et al., 2010; Dube et al., 1999; Hansen & Arntzen, 2013, May; Horsley et al., 2014; Perez, Endemann, Pessôa, & Tomanari, 2015; Pessoa, Huziwara, Perez, Endemann, & Tomanari, 2009; Steingrimsdottir & Arntzen, 2016; Tomanari et al., 2007), and we shall return to a behavior analytic analysis of this phenomenon in the section on “parameters influencing matching-to-sample performance.” First, however, let us see how looking increases the development of stimulus control.

Stimulus Control

When a response differs in the presence and absence of a given stimulus, the response is said to be under stimulus control (Catania, 2013; Schlinger & Poling, 1998; Urcuioli, 2013). For instance, when my phone rings, it plays the song “Hello” by Adel; and I pick it up and say “hello, how can I help you.” I do not pick up my phone to say hello under other circumstances. On the other hand, when my phone plays the auditory stimulus of an old-fashioned typewriter, I pick it up to check my work email—I usually do not check my work email, when I do not hear this sound. Hence, my behavior differs in the presence and absence of the auditory stimulus of Adele singing “Hello” and in the presence and absence of the auditory stimulus of an old-fashioned typewriter. The behaviors of picking up the phone in order to introduce myself and picking up the phone to check my work email have thus come under control of two different stimuli—the auditory stimulus of Adele singing “Hello” and the auditory stimulus of an old-fashioned typewriter, respectively. Such stimuli that evoke specific responses are also referred to as discriminative stimuli. In situations where more than one discriminative stimulus evokes a specific response, we speak of stimuli that are functionally equivalent.

Stimulus Equivalence

Again, research on conceptual and symbolic behavior is the study of how stimuli become equivalent to each other, that is, how different stimuli come to elicit or evoke the same response. In that regard, stimulus equivalence is often referred to as stimulus substitutability (Green & Saunders, 1998). The phenomenon of stimulus equivalence, or stimulus substitutability, has been subject of interest as early as in the 1920’s and 1930’s (e.g., Cofer & Foley Jr, 1942; Hull, 1939; Razran, 1939; Riess, 1940; Watson, 1930;

Weiss, 1925). For instance, Riess (1940, as cited in Keller & Schoenfeld, 1950), carried out an experimental study on mediated generalization, successfully demonstrating the phenomenon of stimulus equivalence. Riess (1940) conditioned the galvanic skin reflex to a linguistic symbol, a word, and then examined whether generalization to two other linguistic symbols, or words, had taken place. One of the words was a synonym to the first word and another word was a homonym of the first word. Results indicated that the conditioned words (i.e., style, freeze, surf, and urn) gained, on average, 346% in the magnitude of the skin response. The synonyms (i.e., fashion, chill, wave, and vase) gained 141% through generalization and the homonyms (i.e., stile, frieze, and earn) gained 94.5%. The generalization of the homonym words demonstrated basic stimulus generalization, whereas the synonyms exemplified “mediated generalization based upon the previous training which produced the 'meaning' equivalence of these two words” (Keller & Schoenfeld, 1950, p. 190).

In other words, Riess’ (1940) experiment demonstrated the emergence of new stimulus-response relations, which had not been directly taught. Thus far, the behavioral research establishment had been quite successful in demonstrating experimental control of behaviors that were directly taught (e.g., paired associates learning). They had also been successful in demonstration "mediated generalization." However, a successful description of this generative phenomenon had not yet been accounted for in a behavior analytic observable and orderly way. Fortunately, and out of curiosity, Sidman (1971) and colleagues began experimenting with the auditory-visual matching-to-sample paradigm in order to teach a developmentally disabled boy reading comprehension—as the experiment turned out successfully, Sidman knew that he had stumbled upon a groundbreaking teaching method.

Sidman's equivalence relation. In short, Sidman (1971) demonstrated that reading comprehension could be taught indirectly to an intellectually disabled boy, who were incapable of reading printed words vocally or with comprehension. Instead, he could match vocalized words to pictures and he could name the pictures. Excitingly, after having taught the disabled boy to match 20 vocalized words to 20 printed words, he was now able to read with comprehension—matching the 20 printed words to 20 images and saying the 20 printed words out loud; the boy had learned an additional 40 relations without direct teach. Indeed, the method demonstrated that new cognitive performances could be generated without having to teach each new piece of information separately.

Notwithstanding successful application of this generative teaching method, it took some years before Sidman and his colleagues were able to string together the relationships between the mathematical and behavioral definitions of equivalence relations—elementary textbooks by Polis, Beard (1973) and Scandura (1971) were studied extensively—both the first and the extended second definition of "the equivalence relation" were first published in two different journals in 1982 (Sidman et al., 1982; Sidman & Tailby, 1982). As for Sidman and Tailby's (1982) extended definition, an equivalence relation must be reflexive, symmetric, and transitive. *Reflexivity* requires that a stimulus is related to itself (i.e., *a*R*a*); *symmetry* requires that two stimuli demonstrate a bidirectional relationship (i.e., if *a*R*b*, then *b*R*a*); *transitivity* requires that two stimuli are related through a third stimulus (i.e., if *a*R*b* and *b*R*c*, the *a*R*c*); and *"global equivalence"* requires a combined test for *symmetry* and *transitivity* (i.e., if *b*R*a* and *c*R*b*, then *c*R*a*) (e.g., see Sidman, Willson-Morris, & Kirk, 1986). The capital "R" means that the two lower case and italicized letters (i.e., symbols for stimuli) are related to each other. Apart from presenting definitions of equivalence relations, another "objective of these papers was to show that behavior-analytic principles and techniques of experimentation could be brought to bear on matters that many considered to require a cognitivist orientation" (Sidman, 1994, p. 119).

To illustrate this innovative and generative teaching method, inherent in a matching-to-sample paradigm (this paradigm will be thoroughly introduced in a following section), a Norwegian child is taught to relate the written Norwegian word "Elg" to a picture of a moose and to relate the picture of a moose to the written English word "Moose" (see Figure 1). In other words, the child is taught two relations via a direct teach and, as a result, four new relations emerge; two symmetric relations, one transitive relation, and a combined relation for symmetry and transitivity —also referred to as global equivalence (e.g., Sidman et al., 1986).

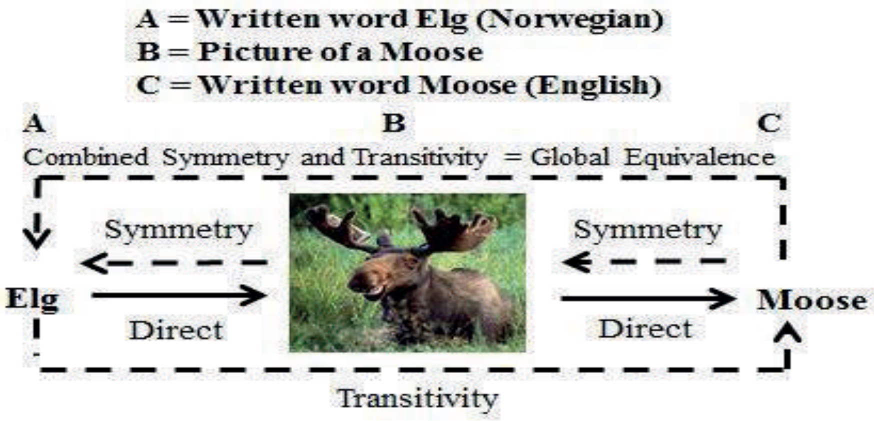


Figure 1. A simplified example of teaching reading comprehension in a matching-to-sample paradigm; two relations are taught directly and, as a result, four relations emerge.

The naming theory. There are, of course, alternative behavioral accounts to the concept of emergence in language development. One account is *The Naming Theory* (Horne & Lowe, 1996), which builds on the idea of verbal naming—it requires (a) listener behavior (i.e., able to respond to others utterances), (b) echoic behavior (i.e., the ability to hear yourself talking), and tacting abilities (i.e., the ability to name objects and events) (e.g., Skinner, 1957). If one possesses these abilities, then (1) listener behavior makes an individual able to hear a spoken word, (2) echoing the word to one self, and (3) tacting the word in the presence of the relevant object, completing the naming circle. In Horne and Lowes own words:

> "We identify naming as the basic unit of verbal behavior, describe the conditions under which it is learned, and outline its crucial role in the development of stimulus classes and, hence, of symbolic behavior" (Horne & Lowe, 1996, p. 185).

The relational frame theory. Another account of generative language development, or equivalence performance, is the *Relational Frame Theory* (RFT) (e.g., Hayes, Barnes-Holmes, & Roche, 2001). The theory is based on "*arbitrary applicable relational responding*," established as a function of *multiple exemplar training* (e.g., Hayes, 1991). In other words, relations between stimuli, *relational frames*, also recognized as generalized operants, emerge as a function of multiple exemplar training; the training of a number of similar stimuli to a number of similar responses. Derived relational frames involve (1) *mutual entailment* (i.e., if A = B, then B = A), (2) combinatorial entailment (i.e., if A = B, and B = C, then A = C), and (3) *transformation of functions* (i.e., if A = B, and A = C, then C = A). In plain English, RFT is:

> "a behavior-analytic account of human language and cognition. It is fundamentally similar to Skinner's account, and is distinct from most cognitive and linguistic

> approaches to language, in that ‘‘it approaches verbal events as activities not products’’ (Hayes, Fox, Gifford, Wilson, & Barnes-Holmes, 2001, p. 22).

With due respect to these alternative approaches to generative language development and cognition, the current project benefits and builds upon the behaviorally systematic and unitary analysis, embedded in the stimulus equivalence paradigm—the emergence of new environment-behavior relations, based on directly taught stimulus-response relations, that is, conditional discrimination training (e.g., Sidman & Tailby, 1982). We begin by introducing a prerequisite, namely, simple discrimination.

Simple discrimination. “A discriminative stimulus (S^D) evokes a response because in the past that kind of response has been more successful in the presence of that stimulus than in its absence” (Schlinger & Poling, 1998, pp. 131–132). Using the example above, the discriminative stimulus (S^D, or S^+), hearing Adele sing “Hello,” evokes the responses of answering my phone by saying “hello,” as this response has been reinforced under similar circumstances. On the contrary, if I answer “hello” when my phone sounds the auditory stimulus of an old-fashioned typewriter, then I will not receive reinforcement, as no one will answer me back. That is, the auditory stimulus of the old-fashioned typewrite will come to serve as an S^{delta} (S^{Δ}, or S^-) for answering my phone by saying “hello” (see Figure 2). An S^{Δ} is a stimulus that suppresses a certain response, as this type of response has been extinguished (i.e., not received reinforcement) in the presence of that stimulus and reinforced in its absence (Schlinger & Poling, 1998).

Figure 2. Illustrated is a diagram of two three-term contingencies: (a) Between an S^D (i.e., phone ringing), a response (i.e., "Hello"), and a reinforcing consequence, S^R ("Hi Steff"); and (b) between an S^Δ, a response (i.e., "Hello"), and a no reinforcing consequence—silence (i.e., extinction).

Thus, a simple discrimination is based on the familiar three-term relation, namely, the analytic unit considered when the two-term reinforcement contingency, response-stimulus relation, comes under control of an environmental discriminative stimulus (Sidman, 1986). The three-term contingency is typically abbreviated S^D: R – S^R. As already touched upon, the two-term contingency, the relation between a behavior and its consequence (i.e., response-stimulus relation) is referred to as the basic unit of analysis and description of operant behavior (as cited in Sidman, 1986; Skinner, 1935, 1938). Making a U-turn, when units of analysis become more complex than basic response-stimulus relations and simple discriminations, that is, under control of yet other environmental stimuli, we refer to such relations as conditional discriminations (e.g., McIlvane, 2013; Urcuioli, 2013).

Conditional discrimination. When a three-term contingency comes under the control of an additional environmental stimulus, we have a four-term contingency—also referred to as a conditional relation. According to Sidman, "the four-term contingency is the fundamental unit of what we might call conditional, instructional, or, as we shall see, contextual stimulus control" (Sidman, 1986, p. 225). A conditional stimulus is typically denoted S^C and the four-term contingency is thus abbreviated S^C: (S^D: – R – S^R). With conditional relations, we often speak of “if, then” relations (e.g., Catania, 2013; Sidman, 1986; Sidman & Tailby, 1982). For instance, if you ask someone a question that can be answered with a “yes” or “no” head movement, then the direction of the head movement for these gestures is dependent on whether you are from either Bulgaria or another country in the world; in Bulgaria, a head nod means no and a head shake means yes—the exact opposite way of gesturing “yes” and “no” than what is customary in most other countries in the world; how we respond to the discriminative stimuli “yes” and “no,” depends on the conditional environmental stimulus, namely, country of living.

In conditional discrimination training, a simple discrimination is trained to come under the control of an additional stimulus—the conditional stimulus (S^C). For example, in Figure 3, a Norwegian child is taught conditional relations that are involved in language comprehension by using the matching-to-sample paradigm (e.g., Sidman & Tailby, 1982): In the presence of the Norwegian word “elg,” (S^C), a child is reinforced (S^R) for selecting (R) the picture of a moose (S^D) but not for selecting (R) the picture of a troll (S^Δ); and in the presence of the picture of a moose (S^C), the child is reinforced (S^R) for selecting (R) the English written word “moose” (S^D) but not for selecting (R) the English written word “troll” (S^Δ).

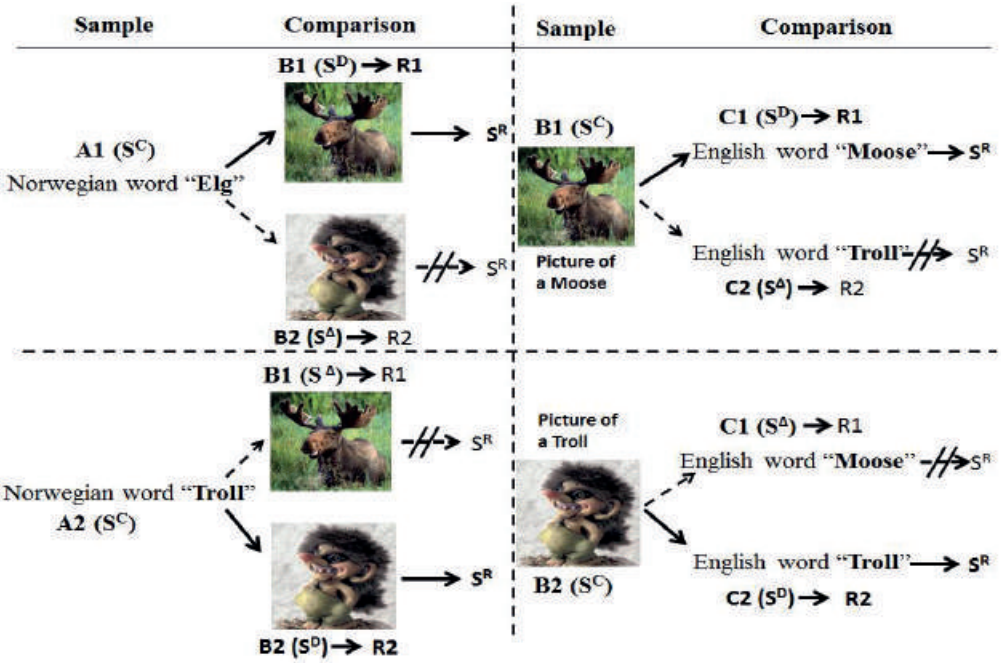


Figure 3. A Norwegian child is taught conditional relations that are involved in reading and language comprehension by using the matching-to-sample paradigm. In a linear series format, A1 is trained to B1, and not B2; and B1 is then trained to C1, and not C2. Concurrently, A2 is trained to B2, and not B1; and then B2 is trained to C2, and not C1.

Still, it is necessary to verify that the taught relations are conditional relations and not just three-term simple discriminations (i.e., relations in which the child is reinforced (S^R) for selecting (R) the discriminative stimuli (S^Ds), namely, the picture of a moose and the written word "moose," regardless of the conditional stimuli (S^Cs), that is, the written word "elg" and the picture of a moose, respectively). Hence, it is important to teach conditional relations in which the roles of the S^Ds and S^Δs are switched; in the presence of the Norwegian word "troll," (S^C), the child is reinforced (S^R) for selecting (R) the picture of a troll (S^D) but not for selecting (R) the picture of a moose (S^Δ); and in the presence of the picture of a troll (S^C), the child is reinforced (S^R) for selecting (R) the English written word "troll" (S^D) but not for selecting (R) the English written word "moose" (S^Δ).

Conditional discrimination procedures vary with regards to degree of complexity. The less complex procedures involve *identity-matching-to-sample* with *familiar stimuli* (e.g., Figure 4), whereas more complex procedures use *arbitrary-matching-to-sample* with either *familiar stimuli* (e.g., Figure 3) or *abstract stimuli*.

Identity matching-to-sample. In identity matching-to-sample, the correct (or positive) comparison stimulus is physically identical to the sample stimulus and, simultaneously, the incorrect (or negative) comparison stimulus is physically different (McIlvane, 2013). Figure 4 exemplifies this arrangement; the comparison stimulus of a red triangle is physically similar to the sample stimulus of a red triangle—picking the identical comparison stimulus results in a reinforcing consequence.

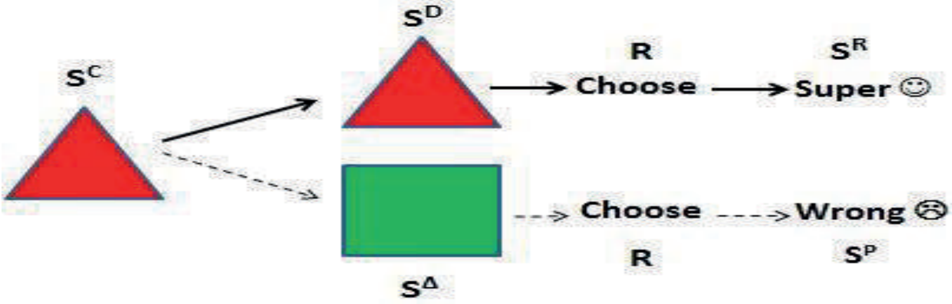


Figure 4. An illustration of a typical identity-matching trial—the correct comparison stimulus is physically identical to the sample stimulus.

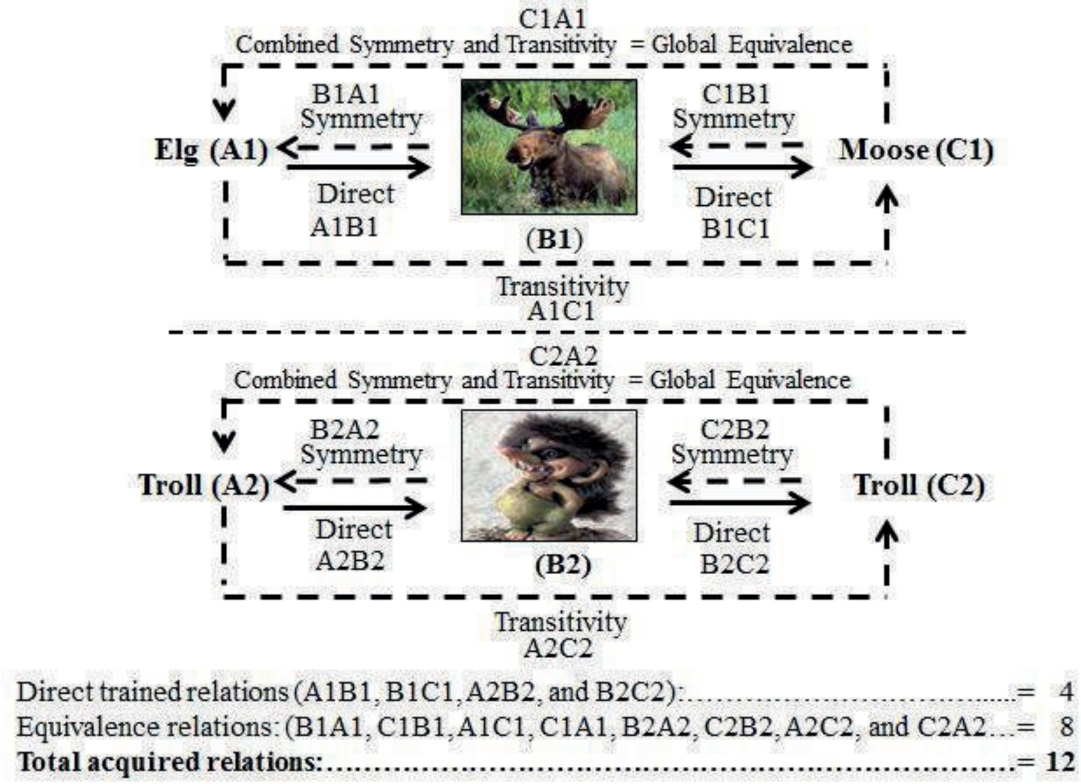
Arbitrary matching-to-sample. On the contrary, arbitrary matching-to-sample involves conditional discriminations, where sample and comparison stimuli are physically different from each other (McIlvane, 2013). Figures 1 and 3 are both examples of arbitrary matching-to-sample—in both examples, the sample stimuli (i.e., picture words of “moose” or “troll,”) are physically different from the comparison stimuli (i.e., pictures of a moose or a troll); choosing the correct comparison stimulus results typically in a generalized reinforcing consequence (e.g., “Super,” or “Great”), as noted in the following.

Familiar stimuli. The familiarity or meaningfulness of a stimulus (i.e., picture, word, or action) can be described by its dictionary defining features and related characteristics (e.g., Fields, Arntzen, Nartey, & Eilifsen, 2012). For instance, smart phones are full of emoji’s (i.e., familiar figures that have acquired emotional characteristics), as the “happy face,” “sad face,” “sleeping face,” or ““thinking face,” and we apply them to text messages in order to clarify intentions and accompanying emotional states.

Abstract stimuli. Meaningless and unfamiliar stimuli are hard to characterize and label, usually because we have no prior history with them. For example, Arabic and Chinese letters are in most cases meaningless to a Scandinavian person who has never been introduced to signs or letters from a second language. In conditional discrimination training, abstract stimuli typically serve to isolate “current discrimination learning processes from prior learning about those stimuli” (McIlvane, 2013, p. 134).

Reinforcement in matching-to-sample arrangements. Reinforcement in conditional discrimination procedures is typically delivered in the form of generalized reinforcers (i.e., social praise, points in a token economy that are exchangeable for something highly preferred), tangibles (i.e., preferable toys to play with), or primary reinforcement in the form of food (e.g., Sidman, 1992a).

Matching-to-sample performance. (Yerkes) originally created the matching-to-sample paradigm to the experimental study of “the mind of a gorilla” (1928, as in Sidman, 1994). Interestingly, the procedure has been more successful with human animals and, furthermore, the thrilling advantage of this conditional discrimination method is, of course, the generation of equivalence relations—an exponential number of additional relations that have not been subject to a direct reinforcement contingency. In that regard, Sidman notes that " positive tests for equivalence permit us to say that the original conditional-discrimination procedure had generated true matching-to-sample performances" (Sidman, 1994, p. 124). This point is very important, as it suggests that we can only speak of matching-to-sample performance when an individual, based on the establishment of conditional discriminations, also demonstrate equivalence class formation—responding in accordance with reflexivity, symmetry, transitivity, as well as a combined test for symmetry and transitivity, that is, global equivalence (Sidman & Tailby, 1982). Such features make the model outstanding in the behavior analytic study of complex behaviors, as they illuminate functional relations in areas that mainstream psychology refer to as language development and cognition (Sidman, 1994).



Direct trained relations (A1B1, B1C1, A2B2, and B2C2):.........................= 4
Equivalence relations: (B1A1, C1B1, A1C1, C1A1, B2A2, C2B2, A2C2, and C2A2...= 8
Total acquired relations:..**= 12**

Figure 5. The four solid arrows represent conditional relations that were taught by explicit reinforcement procedures and the eight broken arrows represent 12 emergent relations.

In a figure similar to Figure 5, Sidman demonstrated that 15 directly taught relations had resulted in 60 emerged relations. He presented the figure at The American Psychological Association and argued that:

> “Cognitivists in many areas—Psychology, Education, Child Development, Linguistics, and Artificial Intelligence, among others—characteristically postulate mental structures to account for complex human behavior. Here […] was an incredible structure—three six-member classes; 15 directly taught conditional relations giving rise to 60 new relations that the subjects had not been explicitly taught. The structure, however, consisted not of postulated mental elements but of

> directly observable relations among elements of the environment. […] It is becoming clear that stimulus equivalence and the cognitive phenomena for which it serves as a basic paradigm can be shown to represent environmental structures. Those structures can be created, rearranged and combined, broken down, and prevented from forming by arranging relations among elements of the environment. The evidence is beginning to place mental structures right back where they started in the first place—in environmental structures" (Sidman, 1994, p. 265).

Parameters Influencing Matching-to-Sample Performance

The following will provide a brief review of some of the variables said to influence the establishment of conditional discriminations and equivalence class formation (for an extended review, please see Arntzen, 2012).

Response requirement to sample stimulus and the observing response. Successful establishment of conditional discriminations is highly dependent on exposure to the conditional stimulus (S^C), the sample stimulus—accuracy has been shown to increase with the requirement of a response to sample stimulus, usually in the form of a mouse click or by using the index finger on a touch screen (e.g., Arntzen, 2012; Arntzen, Braaten, Lian, & Eilifsen, 2011; McIlvane, 2013). In some studies, a response to "feedback" is required as well (Fields, Adams, Verhave, & Newman, 1990). These accompanying, manual, responses are intended to function as observing responses, defined as "acquired environment-behavior relations whose primary function is to affect the sensing of stimuli" (Donahoe & Palmer, 1994, p. 156). Looking at the observing response as a second stage to attending, Skinner wrote that:

> "To attend to something […] is to respond to it in such a way that subsequent behavior is more likely to be reinforced. The precurrent behavior may be learned or

> unlearned. When we turn our eyes toward an object and focus upon it, or sniff an odor, or move a liquid about on the tongue, or slide our fingers over a surface, we make a stimulus more effective. There are two stages: (1) attending to a given state of affairs, and (2) responding to it in some other way. In the normal course of events the reinforcement of the second stage strengthens the first" (Skinner, 1968, p. 122).

In short, the observing response is strengthened by the access to discriminative stimuli and, correspondingly, attending is conditionally strengthened as a prerequisite to the observing response. As such, the initial requirement of a response to sample stimulus will begin to happen voluntarily as stimulus control develops between observing and the four-tem contingency.

The ocular observing response. When our eyes look at and attend to visual stimuli, we engage in ocular observing responses. If the stimuli involved in a conditional discrimination procedure are meaningless (i.e., to me, French picture words would be meaningless as I do not comprehend the written or spoken French), then looking at and attending to the comparison stimuli of meaningless French picture words (e.g., tropic animal names of "snakes" and "monkey," that is, "serpent," and "singe," respectively) will, initially, result in ocular observing responses that do not reliably produce reinforcing stimuli, as one do not yet know which one of the words that relate to the sample pictures of either snake or monkey. Hence, contingent on the sample-picture (e.g., monkey), the choice of one comparison-word will serve as an S^{Δ} ("serpent") for the production of a reinforcing stimulus—instead, the incorrect choice will produce a programmed consequence in the form of the written word "Wrong")—and the choice of the other comparison-word will serve as an S^{D} (singe) for a reinforcer-producing stimulus (i.e., "Correct"). Initially, one would engage in ocular observing responses to both French words (i.e., both S^Ds and S^Δs) equally often, but as stimulus control strengthens, and the French words are reliably chosen and comprehended, ocular observing response to correct comparison stimuli (S^Ds) occur more often and ocular observing responses to incorrect comparison stimuli (S^Δs) occur less often; ocular observing responses to correct comparison stimuli have been conditionally reinforced. Research verifies such contingencies (e.g., Dube et al., 2006; Huziwara, Silva, Perez, & Tomanari, 2015; Perez et al., 2015).

Training structures. Despite Sidman's "bag" analogy—which suggested that an equivalence relation could be thought of as a bag full of all the member-pairs that a contingency specified, without regard to spatial or temporal relationships (e.g., Sidman, 1994, 2000)—it is far from all conditional discrimination procedures that generate true matching-to-sample performance (e.g., Arntzen, Grondahl, & Eilifsen, 2010; Arntzen & Hansen, 2011; Fields, Hobbie, Adams, & Reeve, 1999; Saunders, Chaney, & Marquis, 2005). In other words, training directionality and number of nodes in the LS training structure have been shown to affect equivalence class formation differentially. In order to fully appreciate the differential effects of training structures, it is important to understand the meaning of *singles and nodes*.

Singles and nodes. Stimuli in a class that stand in relation to at least two other stimuli are referred to as nodes. In contrast, stimuli in a class that relate to only one other stimulus are defined as singles (Fields & Verhave, 1987). As we will learn in the following, singles and nodes serve various behavioral functions during conditional discrimination training, depending on training structure used.

Training Directionality. In studies on conditional discrimination training and testing for equivalence class formation, which is matching-to-sample performance, three structures lent themselves useful. These are the linear series (LS), many-to-one (MTO) or comparison as node (CaN), and one-to-many (OTM) or sample as node (SaN) training structures (see Figure 6). In the MTO structure, two or more sample stimuli are related to one comparison stimulus (e.g., BA and CA). As for the OTM structure, one sample stimulus is trained to two or more comparison stimuli (e.g., AB and AC). In the LS structure, however, the nodal stimulus switch function; when training the AB relation, the B stimulus serves as comparison stimulus; and when training the BC relation, the B stimulus severs as sample stimulus (e.g., Arntzen, 2012; Saunders, Saunders, Williams, & Spradlin, 1993; Sidman, 1994).

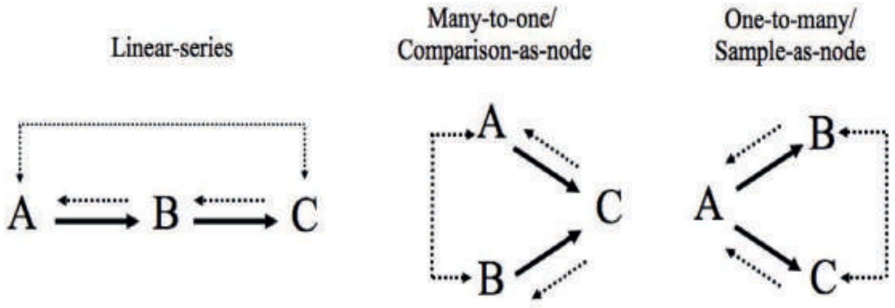


Figure 6. The three most common training structures in matching-to-sample tasks. From left: Linear series (LS), many-to-one (MTO), and one-to-many (OTM) (as illustrated in Arntzen, 2012).

In a high number of studies, the MTO structure has demonstrated to be most effective (e.g., Fields et al., 1999; Hove, 2003; Saunders et al., 1993; Saunders, Chaney, & Marquis, 2005; Saunders, Drake, & Spradlin, 1999) whereas other investigations lean towards best results in the OTM structure (e.g., Arntzen & Holth, 1997; Arntzen & Holth, 2000). Regardless, the LS training structure has consistently been shown to yield the lowest outcome on tests for equivalence class formation—especially when training arrangements have followed a simultaneous protocol (Imam, 2006), as shall be discussed in a section further ahead. First, it seems appropriate to review Saunders and Greens' (1999) discrimination analysis of training structures' effects on equivalence class formation.

Discrimination analysis. Saunders and Greens' (1999) proposal of training structures' effects on equivalence class formation was embedded in the analysis of successive simple discriminations and simultaneous simple discriminations, contained in the conditional discrimination arrangement and testing within the different training structures. Consistent with their analysis, the MTO structure presents all the simple discriminations that are required for repeated positive results on tests for equivalence class formation. In contrast, the OTM structure does not present all the simple discriminations that are necessary for consistent positive outcomes on tests for both symmetry and equivalence. With regards to the LS structure, all simple discriminations required for consistent positive outcomes on tests for symmetry are presented, however, not for tests on transitivity or equivalence. As a result, when number of classes and members increase, the grand total of simple discriminations not presented in each one of the three training structures increase exponentially. What follows is that the MTO structure, having presented all the necessary component simple discriminations during training, should consistently yield better outcomes on tests for emergent relations. The intriguing proposal has been tested out in a number of studies and thus far exponential results, as class size increases, have not yielded the proposed difference (e.g., Arntzen et al., 2010; Arntzen & Hansen, 2011).

Nodal number effect in the linear series structure. Fields and Verhave (1987) looked at response strength in the LS training structure as a function of nodal stimuli. Attempting to define “nodal distance” or “associative distance,” they established that stimuli with a lower nodal number in a stimulus class also had stronger relationship to each other than stimuli with a higher nodal number (e.g., Fields, Reeve, Varelas, Rosen, & Belanich, 1997; Fields & Watanabe-Rose, 2008). In contrast, other studies propose that response strength in the LS training structure has nothing to do with the effects of hypothetical constructs as “nodal distance” or “associative distance.” Behavior analytically, it has been suggested that a decrease in response strength, as number of nodes increase, is a result of a different learning history, that is, the reinforcement-contingency (Imam, 2006).

Training and testing protocols. Different types of training and testing protocols are employed in the establishment of equivalence relations, namely, the simultaneous protocol (SP), the simple-to-complex (STC) protocol, and the complex-to-simple (CTS) protocol (e.g., Imam, 2006). Applying the simultaneous protocol, all baseline conditional relations are trained first before testing formation of equivalence relations (i.e., symmetry, transitivity, and global equivalence). Implementing the STC protocol, one baseline conditional relation (e.g., AB) is established and tested for symmetry (i.e., BA) before a new baseline conditional relation is established (e.g., BC) and tested for symmetry (i.e., CB). Then transitivity (i.e., AC) and global equivalence (i.e., CA) relations are tested. As for the CTS protocol, contrary to the STC protocol, related baseline conditional relations are trained first (e.g., AB and BC) and then tested for global equivalence (i.e., CA). If the equivalence test fails, tests for symmetry (i.e., BA and then CB) and transitivity (i.e., AC) are applied. Test for intact equivalence relations is then repeated and, when passed, new relations are added. Noteworthy, research has consistently produced lowest equivalence yields, when the simultaneous protocol has been implemented (e.g., Fields, Landon-Jimenez, Buffington, & Adams, 1995; Imam, 2006).

Serialized versus concurrent simultaneous training arrangements. Serialized and concurrent training arrangements are sub-protocols of the simultaneous training and testing protocol, in which all conditional relations are trained before testing for baseline-trained, symmetry, and equivalence relations simultaneously (Arntzen, Halstadtro, Bjerke, Wittner, & Kristiansen, 2014; Eilifsen & Arntzen, 2014). In a serialized training format, for example, AB conditional relations are trained and mastered before training the BC conditional relations. On the contrary, in a concurrent training format, all conditional relations, for example AB and BC conditional relations, are trained in a simultaneous and mixed format before testing baseline-trained, symmetry, and equivalence relations simultaneously. Eilifsen and Arntzen showed that participants performed better when they were prepared in a serialized training format (i.e., 50% responded in accordance with stimulus equivalence), compared with participants with conditional relations trained in a concurrent arrangement (i.e., 10% responded in accordance with stimulus equivalence). Furthermore, one of the studies (i.e., Arntzen, Halstadtro, et al., 2014) suggested better follow-up performance on test for equivalence class formation, when the participant was trained in a serialized training format. Hence, it was proposed that a serialized training format could help explain the relative high performance ratio noted with participants in the LS condition.

Common dependent measures in matching-to-sample tasks. In studies on conditional discrimination training and testing for equivalence class formation, the results section usually report data on dependent measures such as (a) number of trials used during training sessions in order to reach criterion for testing; (b) number and percentage of trials correct during training; (c) number of errors during training (e.g., Arntzen et al., 2010; Arntzen & Hansen, 2011; Arntzen, Nartey, & Fields, 2014; Fields et al., 1999; Pilgrim, Jackson, & Galizio, 2000); (d) reaction times to comparison stimuli for (i) baseline-conditional relations in the last phase of training and the beginning and end of testing, (ii) symmetry relations during the beginning and end of testing, and (iii) transitivity/equivalence relations during the beginning and end of testing (e.g., Arntzen & Holth, 1997; Holth & Arntzen, 2000); (e) speed (i.e., the inverse reaction time) during comparison choice selection in training and testing conditions (e.g., Arntzen, Petursson, Sadeghi, & Eilifsen, 2015; Spencer & Chase, 1996); and (f) pre- and post-sorting performance (e.g., Arntzen, Granmo, & Fields, 2016). A relative new dependent measure to record in conditional discrimination procedures and matching-to-sample performance tasks, however, is eye-movements (Dube et al., 2006; Dube et al., 2010; Steingrimsdottir & Arntzen, 2016).

Eye-movements as additional dependent measures. Continuous advancements in technology—and the engineering of eye-tracking equipment—provide opportunities to record eye-movements during conditional discrimination training and testing for equivalence class formation. As mentioned earlier, Yarbus (1967) established eye-movement topographies as eye-fixation, saccade, and smooth pursuit. Eye-fixations were described as “sensed visual stimuli that are stationary relative to an observer’s head and eyes." Hence, we engage in eye-fixations when our eyes come to a relative halt and look directly at various forms of visual stimuli[1]. As a result, new dependent measures as (1) an eye-fixation and (2) the duration of an eye-fixation, pertinent to behavioral analytic research, can now be recorded in complex visual displays.

In addition to reaction time, therefore, the progressive developments in eye-tracking technology now provide us with the means to record (1) eye-fixation durations to sample stimuli and each of the comparison stimuli in the comparison choice array and, furthermore, (2) number of eye-fixations to sample stimuli and each of the comparison stimuli in the comparison choice array—during conditional discrimination training and testing for equivalence class formation.

The Studies in the Dissertation

Ethical concerns are crucial to research with human participants. Deliberate considerations of such issues are first followed by a short introduction to the studies involved in the dissertation, then leading to a general discussion on new findings.

Ethical Concerns

When conducting research in experimental and applied settings, it is crucial to comply with ethical standards and proposals that are approved by local Human Research Committees (e.g., in Norway, "Regionale Komiteer for Medisinsk og Helsefaglig Forskningsetikk" ("REK")—Regional Committees for Medicine and Health Related Research Ethics) and, if required, local Institutional Review Boards (IRBs). Therefore, the Behavior Analyst Certification Board has lined out guidelines for responsible conduct for behavior analysts and the American Psychological Association has issued "*Ethical*

---

[1] Structural and functional understandings of eye-fixations are considered in Hansen and Arntzen (Hansen & Arntzen, 2015).

Principles of Psychologists and Code of Conduct;" both describe standards of professional conduct and ethical practice (e.g., Cooper, Heron, & Heward, 2007).

Informed consent. After welcoming participants to our experimental research lab, they were introduced to a consent form, which they had to sign before initiating an experiment. Also, they received detailed information about the experimental situation, namely, how we used eye-tracking equipment to acquire knowledge about eye-movements during problem-solving tasks. In addition, we let participants know that a clarifying debriefing, with regards to research question, would be giving upon completion of experiment. Finally, we informed participants that they could leave the experimental situation at any time, of course, without any negative consequences (e.g., Bailey & Burch, 2016).

Do no harm and right to withdraw. When conducting research with student participants, it is crucial to take precautionary measures in order to avoid doing harm. The Behavior Analyst Certification Board defines harm as negative effects or side effects that outweigh positive effects in each specific instance and that, in addition, are directly observable (Cooper et al., 2007). In our experiments, we made sure to let participants know that they would not be introduced to harmful situations, in form of physical or emotional pain. If, for any reason, participation felt unpleasant, participants were told to let the experimenter know right away.

Privacy. Following an experiment, behavior analysts and researchers do not disclose confidential information about participants. Neither do researchers talk about experimental situations, concerning the performance of specific participants, in such a way that the identity of the specific participant is revealed. Before initiating one of our experiments, for instance, participants were introduced to a specific test-situation which had already received a "participant number." These participant numbers were never paired up with the exact participant, hence, results could never reveal a participant's identity (Bailey & Burch, 2016).

Role of the Experimenter. During experimental situations, investigators or experimenters interfere with participants as little as possible and only in ways that are necessary, with respect to a particular research design. For instance, as experimenters, we told participants that we would enter the experimental facility every 25th minute in order to do a quick recalibration of the eye-tracking equipment; we would say only a few words, all related to places to look at the screen during calibration. Then we would, discretely, exit the experimental room again. Of course, as experimenters, we would always make sure that participants were feeling well during all parts of the experiment.

General principles. In addition, The American Psychological Association (Association, 2002) has issued some general principles with the intent to inspire behavior analysts and psychologist to the highest ethical standards of their profession. The five general principles are (1) Beneficence and Nonmaleficence, (2) Fidelity and Responsibility, (3) Integrity, (4) Justice, and (5) Respect for People's Rights and Dignity.

Beneficence and nonmaleficence. Psychologists and behavior analysts make every effort to help those around them and strive to do no harm. Professionally, they try to defend the well-being and privileges of people they interact with professionally. If encounters arise among behavior analysts, they try to solve these in mature way. Furthermore Psychologists strive to be aware of the possible effect of their own physical and mental health on their ability to help those with whom they work.

Fidelity and responsibility. Behavior Analysts and psychologists strive to develop honest and trustworthy relationships with colleagues. Moreover, they are conscious about their relationships of trust with those with whom they work. They are aware of their professional and scientific duties to the community. Also, they maintain professional ideals and guide themselves by the codes of conduct.

Integrity. Behavior analysts and psychologists seek to uphold truth, morality and honesty in science and research. Moreover, they do not steal, cheat, or involve themselves in related activities.

Justice. Behavior analysts and psychologists are aware that equality and fairness permit all individuals of the right to use and take advantage of the offerings from behavior analysts and psychologists. Furthermore, behavior analysts and psychologists are careful not to overlook unfair practices.

Respect for people's rights and dignity. Behavior analysts and psychologists respect the self-worth and dignity of all human beings. Also, they acknowledge people's civil privileges in the form of privacy, confidentiality, and self-determination.

About the Studies

In the first study, our goal was three-fold; namely, (1) to introduce eye-tracking equipment and appropriate fixation-detection algorithms to the behavior analytic community thru an extensive literature review; (2) to acquire a behavior analytic understanding of eye-movements, specifically, how eye-fixations relate to attending, observing, and perceiving (i.e., stimulus control); and (3) to provide behavioral enthusiasts with dependent fixation measures, readily identifiable with eye-tracking equipment and appropriate identification algorithms. First, we learned that eye-tracking technology is increasing in popularity, as it reveals where a person is looking. Therefore, a growing number of enterprises and organizations are taking advantage of this tool. Dependent on aim of use, many types of eye-trackers and algorithms have been developed—for example, if using eye-tracking equipment while involved in a conditional discrimination procedure, presented on a computer screen, then literature suggests one type of eye-fixation detection algorithm, which is the detection-threshold identification (I-DT) algorithm—this algorithm is appropriate for such an activity, as eyes move relatively slow and operate within a limited area.

As for our second goal, we looked through solid scientific research papers in order to gain a sound behavior analytic understanding of eye-movements and the functional relationships that exist between fixating, attending, and observing—with the ultimate goal of finding an answer to the question: when can we say that a person has emitted an ocular observing response to a visual discriminative stimulus? On our journey to find a solid behavior analytic answer to this question, we learned that ocular observing responses occur with and without clear-cut eye fixation and, also, that ocular observing responses vary across behaviors, settings, and individuals. Arriving at an answer, or guiding rule, for our upcoming experimental studies, we established 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" (Hansen & Arntzen, 2015, p. 16). Regarding our third goal, we reviewed an extensive amount of literature in order to learn more about dependent fixation measures—we identified appropriate dependent measures as rate, number, proportion, and pattern.

In study 2, we used eye-tracking equipment to record eye-movements during matching-to-sample performance, as dependent fixation measures (i.e., duration and rate) had yet to be accounted for as a function of training directionality. With nine university-college students, introduced in a within-subjects design, we systematically replicated previous research on the differential effects of training structures; potential six 3-member equivalence classes were established with MTO, OTM, and LS training structures—conditional relations were introduced in a serialized arrangement, using a simultaneous training and testing protocol. Results showed that one of three participants, one in each preparation, formed equivalence classes. Furthermore, participants who formed equivalence classes revealed longer fixation times to sample stimuli and shorter fixation durations to comparison stimuli. On the other hand, participants recorded longer fixation durations and higher fixation rates to correct comparison stimuli, regardless of demonstrating equivalence relations—a finding that has implications with respect to eye-movements predictive value on delayed emergence of stimulus equivalence class formation.

In study 3, we systematically replicated study 2, that is, we explored the differential outcomes in fixation times and fixation rates as a function of training directionality. In contrast to study 2, we employed a group-design; thirty participants were randomly assigned to one of the three training structures, MTO, OTM, and LS, and conditional relations were, this time, presented concurrently, using a simultaneous training and testing protocol—we suspected that the more challenging, concurrent training format would aid us in our attempts to identify molecular differences among the three training structures. Also, to moderate time in the experimental situation, we reduced number of classes; this time we hoped to form five 3-member equivalence classes. Furthermore, whereas we in study 2 presented 72 baseline-trained, 36 symmetry, and 36 equivalence trials during emergent relations testing, in study 3 we decided to equalize number of presented test trials, namely, 30 baseline-trained, 30 symmetry, and 30 equivalence trials. Results showed that participants in the MTO group performed best, with seven of ten generating equivalence classes. However, they used notably more trials to complete the training phase. Moreover, results on dependent fixation measures replicated findings from study 2; specifically, (a) participants who formed equivalence classes revealed longer fixation times to sample stimuli during training; (b) participants fixated longer and more often to correct comparison stimuli, regardless of equivalence class formation; and, (c) generally, the MTO condition revealed longer fixation times and higher fixation rates to comparison stimuli, especially in the initial training phase and during testing for equivalence relations.

General Discussion

Introducing eye-tracking technology, the purpose of the three studies in the dissertation was to increase our knowledge about eye-movements and their functional relations to complex human behavior—in order to obtain a conceptual understanding of the dependent eye-movement measure, *fixating*, and its functional relation to behaviors as attending, observing, and perceiving, we reviewed scientific literature on eye-fixations and observing responses. Also, taking advantage of such knowledge, and to broaden our behavior analytic understanding of complex human behavior, we introduced experiments in which we recorded and analyzed dependent eye-movement measures, as fixation duration and fixation rate, during differentially prepared matching-to-sample tasks.

Conceptually and functionally; we learned (1) that eye-fixations are defined as "sensed visual stimuli that are stationary relative to an observer's head and eyes;" (2) that an individual has fixated at, attended to, and observed a visual discriminative stimulus—long enough to perceive it—when the visual discriminative stimulus reliably causes that individual to respond in accordance with the experimenter-defined contingencies; (2) that ocular observing responses happen with and without clear-cut eye-fixation; and, furthermore, (3) that ocular observing responses are context-specific, which means that they vary across behaviors, settings, and individuals.

Experimentally, (a) in study 2, we did not find noteworthy difference in the formation of equivalence classes, based on training directionality; (b) in study 3, on the other hand, results showed that participants in the MTO group performed best, with seven of ten generating equivalence classes—they used, however, notably more trials to complete the training phase. Generally, (c) participants who formed equivalence classes, recorded longer fixation times to sample stimuli during training; (d) they fixated longer on incorrect trials; and, most noteworthy, (e) they fixated longer and more often at correct comparison stimuli, irrespective of training directionality and equivalence class formation.

Directionality and Equivalence Class Formation

Abstract stimuli were used in both of our experimental studies (i.e., study 2 and 3), as this would help prevent previous learning from interfering with our planned independent measures—this is in accordance with McIlvane's (2013) notion that abstract stimuli typically serve to isolate "current discrimination learning processes from prior learning about those stimuli" (McIlvane, 2013, p. 134). Nevertheless, the independent variable in study 2, that is, training structure, did not reveal noteworthy differences. Nine participants were presented to one of the three training structures—making three in each condition—and one participant per training structure formed emergent relations. Furthermore, all but one responded in accordance with symmetry. Needless to say, participants in the LS group performed surprisingly well; previous research has shown that the LS condition performs worse than the OTM and MTO conditions, especially when prepared in a serialized training format (e.g., Arntzen, Grondahl, & Eilifsen, 2011). One procedural issue that could improve test performance in the LS condition was the high number of baseline-trained relations that were interspersed among symmetry and equivalence trials; the test conditions consisted of 72 baseline-trained, compared to 36 symmetry and 36 equivalence trials, respectively. In support, previous research findings have shown that extended exposure to intact test trial types aids the establishment of other trial types (e.g., Spradlin, Saunders, & Saunders, 1992). On the other hand, it could also be argued that the MTO and OTM conditions performed worse than expected, as only 30% demonstrated matching-to-sample performance; other studies have demonstrated that equivalence performance ratios for these two structures normally reach about 90% (e.g., Arntzen, 2012; Arntzen & Hansen, 2011; Saunders et al., 2005).

Hence, in order to increase the likelihood of exposing the molecular differences that are theoretically inherent in the three training structures, study 3 systematically replicated study 2; (a) instead of six classes (a bit exhaustive to participants, as they had to sit in the same position over a longer period of time), we introduced five classes; (b) instead of a serialized training format, we presented conditional relations concurrently (i.e., mixed); (c) in contrast to an uneven number of test trial types, study 3 presented an even number, that is, 30 test trials for baseline-trained, symmetry, and equivalence relations, respectively; and (d) instead of a single-subject design, with data decisions relying on visual inspection, study 3 introduced a group design with ten participants in each group—thus, allow for a statistical analysis of findings. Results of such manipulations demonstrated better performance among participants in the MTO condition, with seven of ten participants generating equivalence classes. Furthermore, participants in this group also recorded notably more trials to reach test criterion (i.e., 90% correct in the last test block with no programmed consequences).

Results of study 3 are in accordance with previous findings from our lab (e.g., Arntzen et al., 2010; Arntzen & Hansen, 2011) and, furthermore, support Saunders and Green's (1999) component simple discrimination analysis; as number of classes increases, the number of simple discriminations not presented during training (in OTM and LS but not in MTO) increases exponentially; as with three classes, the number of simple discriminations not presented, in OTM and LS training, is four and with five classes, this number exponentially increases to 25—yet, all component simple discriminations, allowing for successful test performance, are presented in the MTO condition (Saunders & Green, 1999). Moreover, the weak matching-to-sample performances in the LS condition also line up along previous findings (e.g., Arntzen et al., 2010; Arntzen & Hansen, 2011; Saunders et al., 2005).

The fact that the nodal stimulus in the LS condition changes behavioral function (i.e., the B stimulus act as comparison stimulus in AB training and as sample stimulus in

BC training) does not explain the differential results that are obtained in study 2 and 3—both procedures involved 3-member classes, that is, one-node stimulus classes.

Serialized versus Concurrent Trial Arrangements

On the other hand, the introduction of the concurrent trial arrangement (i.e., randomly presented, all conditional relations are trained and tested simultaneously) in study 3—a trial arrangement that is known to decrease the likelihood of forming equivalence relations (e.g., Arntzen, Halstadtro, et al., 2014; Eilifsen & Arntzen, 2014)—could help explain the differential results between study 2 and 3, as this arrangement aids the identification of strengths and weaknesses that are embedded in the MTO, OTM, and LS training structures. For instance, with LS, when establishing conditional relations in a serialized, or sequential, arrangement, one conditional relation is typically established before moving on to the next conditional relation. Such an arrangement allows for development of stimulus control, while the nodal stimulus serves as comparison; and a change in behavioral function (i.e., from serving as comparison to serving as sample stimulus) will thus happen more smoothly—the nodal stimulus will not have to serve two functions simultaneously, as is the case when presented in the concurrent trial arrangement, and the probability of forming equivalence classes is therefore higher.

Dependent Fixation Measures

The systematic manipulations in study 3 resulted in dependent fixation recordings that were similar to previous findings (i.e., study 2); participants who formed equivalence classes recorded longer fixation durations to sample stimuli during establishment of conditional relations and, furthermore, longer durations and rates to correct comparison stimuli—this regardless of demonstrating matching-to-sample performance. Results were most pronounced in the MTO structure, as changes in fixation rates were significant both during training and testing—and only significant during training for participants prepared with the LS structure. General findings on dependent fixation measures replicate Steingrimsdottir and Arntzen (2016) and other researchers (e.g., Huziwara, de Souza, & Tomanari, 2016; Perez et al., 2015; Tomanari et al., 2007). Future studies should also study fixation times and fixation rates as a function of an increase in class size—especially with regards to the LS structure; as number of nodes increases, will participants still show longer fixation durations and higher fixations rates to correct comparison stimuli, both with regards to equivalence class performers and non-equivalence class performers?

Study 2 and 3 indicate longer fixation times and higher fixation rates to correct comparison stimuli (i.e., S^+ stimuli), even among participants who do not form equivalence relations, and this forces us to repeat the question that was asked in study 1: Will significant higher dependent fixation measures to correct comparison stimuli, among participants who do not demonstrate emergent relations after first test, allow us to predict delayed emergence of equivalence relations? The phenomenon of delayed emergence of equivalence class formation has been investigated for the past forty years (e.g., Spradlin, Cotter, & Baxley, 1973; Spradlin et al., 1992) and the research area is experiencing a renaissance (e.g., Lian & Arntzen, 2015). Hence, to answer this question, future studies should extend testing with non-equivalence class performers; based on their dependent fixation measures to comparison stimuli during initial testing, will it be possible to predict who will and who will not demonstrate delayed emergence?

When Does an Equivalence Class Emerge?

Sidman has considered equivalence as (a) a fundamental process and (b) as a result of the reinforcement contingency (Sidman, 2000). Regarding equivalence as fundamental process, Sidman has proposed that equivalence is a fundamental stimulus function, like reinforcement and discrimination, that is, not derivable from other behavioral processes (Sidman, 1992a). On the reinforcement contingency, Sidman notes that it

> "produces two types of outcome: (a) 2-, 3-, 4-, 5-, or n-term units of analysis that are known, respectively, as operant reinforcement, simple discrimination, conditional discrimination, second-order conditional discrimination, and so on; and (b) equivalence relations that consist of ordered pairs of all positive elements that participate in the contingency" (Sidman, 2000, p. 127).

Hence, Sidman suggests that an equivalence class emerges directly from the reinforcement contingency. Our own results support this view, as we have found differential results in dependent eye-fixation measures, as duration and rate, to (a) different training structures, (b) correct and incorrect test trials, (c) correct and incorrect comparison stimuli, and (d) different test trial types (i.e., baseline-trained, symmetry, and equivalence relations).

Limitations to Study 2 and 3

We have identified some common limitations to the experimental studies in the dissertation. The first concerns peripheral vision; a limitation also pointed out by Perez and colleagues (2015). As discussed in study 1, peripheral vision is a methodological concern when eye-tracking technology is implemented in MTS tasks. In study 2, we elaborated on the systematic change in dependent fixation measures, as a function of learning experience. That is to say; as conditional stimulus control develops, there will be less “clear-cut” eye-fixations to correct comparison stimuli. We supported this reasoning by pointing to Skinner (1953), who wrote that “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). In our experimental studies, we have taken this into consideration; when evaluating data, it is necessary to pay attention to changes in dependent measures as a whole; is a proportional change in data identifiable? If so, then justification of experimental findings will be served.

Calibration was another concern of ours; were we able to register precise and valid eye-movements during the course of an experimental situation? To combat this challenge, as explained in the method sections, we arranged for a re-calibration every 25th minute. In addition, and based on our own recommendations from study 2, study 3 introduced TeamViewer (https://www.teamviewer.com/en); it provided us with the opportunity to follow the experiment in real time; every 25th minute notwithstanding, now we could, discretely, arrange for recalibrations when warranted—a solid improvement in validity.

Concluding Remarks

Eye-tracking has provided us with the opportunity to study complex behavior from a new perspective—now we can observe behavior thru the eyes of the individual under study and, as a result, learn more about his or her behavior during the response delay.

Also, in our first experimental study, we asked the question: which one of the dependent measures, fixation time or fixation rate, will serve as a better illuminator of the molecular differences that are embedded in the three training structures? Based on our findings in Study 2 and Study 3, we suggest that fixation rate is the better indicator of molecular differences. As fixation rate is concrete, that is, an all or none measure (i.e., a person fixates inside the stimulus area, and we register a fixation, or else she does not), it is likely that this dependent measure will show more variation as a function of learning. Fixation time will, of course, also vary but not with the same magnitude; it is not an all or none dependent measure, therefore, calculated medians of fixation times will show less variation than is the case with number of trials per training block.

This debate notwithstanding, I think that it is safe to say that our experimental studies have demonstrated that eye-tracking technology and eye-movements add to our understanding of complex human behavior.

References

Arntzen, E. (2012). Training and Testing Parameters In Formation of Stimulus Equivalence: Methodological Issues. *European Journal of Behavior Analysis, 13*, 123–135. Retrieved from http://www.ejoba.org/

Arntzen, E., Braaten, L. F., Lian, T., & Eilifsen, C. (2011). Response to sample requirements in conditional discrimination procedures. *European Journal of Behavior Analysis, 12*, 505–522. Retrieved from http://www.ejoba.org/

Arntzen, E., Granmo, S., & Fields, L. (2016). The Relation Between Sorting Tests and Matching-to-Sample Tests in the Formation of Equivalence Classes. *The Psychological Record*, 1–16.

Arntzen, E., Grondahl, T., & Eilifsen, C. (2010). The effects of different training structures in the establishment of conditional discriminations and subsequent performance on tests for stimulus equivalence. *The Psychological Record, 60*, 437–461. Retrieved from http://opensiuc.lib.siu.edu/tpr/vol60/iss3/4.

Arntzen, E., Grondahl, T., & Eilifsen, C. (2011). The effects of different training structures in the establishment of conditional discriminations and subsequent performance on tests for stimulus equivalence. *The Psychological Record, 60*, 4.

Arntzen, E., Halstadtro, L., Bjerke, E., Wittner, K. J., & Kristiansen, A. (2014). On the sequential and concurrent presentation of trials establishing prerequisites for emergent relations. *The Behavior Analyst Today, 14*, 1.

Arntzen, E., & Hansen, S. (2011). Training Structures and the Formation of Equivalence Classes. *European Journal of Behavior Analysis, 12, 483–503., 12*.

Arntzen, E., & Hansen, S. (2103, April). *Eye-tracking technology as an additional measurement in conditional discrimination procedures.* Paper presented at the Tobii Eye Tracking Conference on Behavioral Research, Boston, MA.

Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record, 47*, 309–320.

Arntzen, E., & Holth, P. (2000). Probability of stimulus equivalence as a function of class size vs. number of classes. *The Psychological Record, 50*, 79–104. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html.

Arntzen, E., Nartey, R., & Fields, L. (2014). Identity and Delay Functions of Meaningful Stimuli: Enhanced Equivalence Class Formation. *The Psychological Record, 64*, 349-360. doi: 10.1007/s40732-014-0066-3

Arntzen, E., Petursson, P. I., Sadeghi, P., & Eilifsen, C. (2015). Equivalence class formation in accuracy or speed conditions: immediate emergence, adduction, and retention. *The Psychological Record, 65*, 141–159.

Association, A. P. (2002). Ethical principles of psychologists and code of conduct. *American psychologist, 57*, 1060–1073.

Bailey, J., & Burch, M. (2016). *Ethics for behavior analysts*: Routledge.

Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols*: New York: Academic Press.

Catania, A. C. (2013). *Learning* (5 ed.). New York, NY: Sloan Publishing.

Cofer, C. N., & Foley Jr, J. P. (1942). Mediated generalization and the interpretation of verbal behavior: I. Prolegomena. *Psychological Review, 49*, 513.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis*. NJ, Merril: Upper Saddle River.

Deacon, T. (1997). The symbolic species: The co-evolution of language and the brain: New York: Norton.

Dodge, R. (1907). An experimental study of visual fixation. *Psychological Monographs: General and Applied, 8*, i–95. doi: 10.1037/h0093042 Retrieved from http://psycnet.apa.org/journals/mon/8/4/i.pdf.

Donahoe, J. W., & Palmer, D. C. (1994). *Learning and complex behavior*: Allyn & Bacon.

Dube, W. V., Balsamo, L. M., Fowler, T. R., Dickson, C. A., Lombard, K. M., & Tomanari, G. Y. (2006). Observing behavior topography in delayed matching to multiple samples. *The Psychological Record, 56*, 6. Retrieved from http://opensiuc.lib.siu.edu/tpr/vol56/iss2/6.

Dube, W. V., Dickson, C. A., Balsamo, L. M., O'Donnell, K. L., Tomanari, G. Y., Farren, K. M. et al. (2010). Observing behavior and atypically restricted stimulus control. *Journal of the experimental analysis of behavior, 94*, 297–313. doi: 10.1901/jeab.2010.94-297

Dube, W. V., Lombard, K. M., Farren, K. M., Flusser, D. S., Balsamo, L. M., & Fowler, T. R. (1999). Eye tracking assessment of stimulus overselectivity in individuals with mental retardation. *Experimental Analysis of Human Behavior Bulletin, 17*, 8–14.

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice* (Vol. 373): Springer.

Eilifsen, C., & Arntzen, E. (2014). Effects of serialized and concurrent training in matching-to-sample procedures. *Submitted for publication*.

Fields, L., Adams, B. J., Verhave, T., & Newman, S. (1990). The effects of nodality on the formation of equivalence classes. *Journal of the Experimental Analysis of behavior, 53*, 345–358.

Fields, L., Arntzen, E., Nartey, R. K., & Eilifsen, C. (2012). Effects of a meaningful, a discriminative, and a meaningless stimulus on equivalence class formation. *Journal of the Experimental Analysis of Behavior, 97*, 163–181.

Fields, L., Hobbie, S. A., Adams, B. J., & Reeve, K. F. (1999). Effects of training directionality and class size on equivalence class formation by adults. *The Psychological Record, 49*, 703–724.

Fields, L., Landon-Jimenez, D. V., Buffington, D. M., & Adams, B. J. (1995). Maintained nodal-distance effects in equivalence classes. *Journal of the Experimental Analysis of behavior, 64*, 129–145.

Fields, L., & Reeve, K. F. (2000). Synthesizing equivalence classes and natural categories from perceptual and relational classes. *Issues in experimental and applied analyses of human behavior*, 59–84.

Fields, L., Reeve, K. F., Varelas, A., Rosen, D., & Belanich, J. (1997). Equivalence class formation using stimulus-pairing and yes-no responding. *The Psychological Record, 47*, 661.

Fields, L., & Verhave, T. (1987). The structure of equivalence classes. *Journal of the experimental analysis of behavior, 48*, 317–332.

Fields, L., & Watanabe-Rose, M. (2008). Nodal structure and the partitioning of equivalence classes. *Journal of the Experimental Analysis of Behavior, 89*, 359–381.

Green, G., & Saunders, R. R. (1998). Stimulus equivalence *Handbook of research methods in human operant behavior* (pp. 229–262): Springer.

Hansen, S., & Arntzen, E. (2013, May). *A review on the use of eye-tracking technology to study observing behavior during Conditional discrimination training.* Paper presented at the ABAI 39th Annual Conference, Minneapolis, MN.

Hansen, S., & Arntzen, E. (2015). Fixating, attending, and observing: a behavior analytic eye-movement analysis. *European Journal of Behavior Analysis, 16*, 229–247. doi: 10.1080/15021149.2015.1090750

Hansen, S., & Arntzen, E. (2016b). Variability in Fixation Measures across Training Structures. *Manuscript submitted for publication*.

Hayes, S. C. (1991). A relational control theory of stimulus equivalence. *Dialogues on Verbal Behavior.* In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on Verbal Behavior*. Reno, Nevada: Context Press.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*: Springer Science & Business Media.

Hayes, S. C., Fox, E. J., Gifford, E. V., Wilson, K. G., & Barnes-Holmes, D. (2001). Derived relational responding as learned behavior. In S. C. Hayes, D. Barnes-

Holmes & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 21–49). New York: Kluwer Academic/Plenum.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*: Oxford University Press.

Holth, P., & Arntzen, E. (2000). Reaction Times and the Emergence of Class Consistent Responding: A Case for Precurrent Behavior? *The Psychological Record, 50*, 305–337.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of behavior, 65*, 185–241.

Horsley, M., Eliot, M., Knight, A., & Reilly, R. (2014). *Current trends in eye tracking research*: Springer.

Hove, O. (2003). Differential probability of equivalence class formation following a one-to-many versus a many-to-one training structure. *The Psychological Record, 53*, 617–634.

Hull, C. L. (1920). Quantitative aspects of evolution of concepts: An experimental study. *Psychological monographs, 28*, i.

Hull, C. L. (1939). The problem of stimulus equivalence in behavior theory. *Psychological Review, 46*, 9.

Huziwara, E. M., de Souza, D. d. G., & Tomanari, G. Y. (2016). Patterns of eye movement in matching-to-sample tasks. *Psicologia: Reflexão e Crítica, 29*, 1-10.

Huziwara, E. M., Silva, M. J. M., Perez, W. F., & Tomanari, G. Y. (2015). Analyzing the stimulus control acquisition in simple discrimination tasks through eye movement. *Psicologia: Reflexão e Crítica, 28*, 603–612.

Imam, A. A. (2006). Experimental control of nodality via equal presentations of control discriminations in different equivalence protocols under speed and no-speed conditions. *Journal of the Experimental Analysis of behavior, 85*, 107–124.

Keller, F. S., & Schoenfeld, W. N. (1950). Principles of Psychology: A systematic text in the science of behavior. . New York: : Appleton-Century-Crofts.

Kirshner, H., & Sidman, M. (1972). Scanning patterns in aphasic patients during matching-to-sample tests. *Neuropsychologia, 10*, 179–184.

Lakoff, G. (1987). Women, fire, and dangerous things: What categories reveal about the mind. *Chicago: University of Chicago*.

Laurence, S., & Margolis, E. (1999). Concepts and cognitive science. *Concepts: core readings*, 3-81.

Lian, T., & Arntzen, E. (2015). *Extended Testing and Delayed Emergence of Derived Relations*. Paper presented at the 41st Annual Convention ABAI. Association for Behavior Analysis International., San Antonio.

McIlvane, W. J. (2013). Simple and complex discrimination learning.

Palmer, D. C. (2002). Psychological essentialism: A review of E. Margolis and S. Laurence (EDS.), Concepts: Core readings. *Journal of the experimental analysis of behavior, 78*, 597-607.

Palmer, D. C. (2010). Behavior Under the Microscope: Increasing the Resolution of Our Experimental Procedures. *The Behavior Analyst, 33*, 37–45. Retrieved from http://www.ncbi.nlm.nih.gov/.

Perez, W. F., Endemann, P., Pessôa, C. V., & Tomanari, G. Y. (2015). Assessing stimulus control in a discrimination task with compound stimuli: Evaluating testing procedures and tracking eye fixations. *The Psychological Record, 65*, 83–88.

Pessoa, C., Huziwara, E. M., Perez, W. F., Endemann, P., & Tomanari, G. Y. (2009). Eye fixations to figures in a four-choice situation with luminance balanced areas: Evaluating practice effects. *Journal of Eye Movement Research, 3*, 1–6. Retrieved from http://www.ufscar.br/.

Pilgrim, C., Jackson, J., & Galizio, M. (2000). Acquisition of arbitrary conditional discriminations by young normally developing children. *Journal of the Experimental Analysis of Behavior, 73*, 177–193.

Polis, A. R., & Beard, E. M. L. (1973). *Fundamental mathematics for elementary teachers: A behavioral objectives approach*. New York: Harper & Row.

Razran, G. H. S. (1939). A quantitative study of meaning by a conditioned salivary technique (semantic conditioning). *Science*.

Riess, B. F. (1940). Semantic conditioning involving the galvanic skin reflex. *Journal of experimental psychology, 26*, 238.

Rosch, E. (1999). Principles of categorization. *Concepts: core readings*, 189-206.

Saunders, K. J., Saunders, R. R., Williams, D. C., & Spradlin, J. E. (1993). An interaction of instructions and training design on stimulus class formation: Extending the analysis of equivalence. *The Psychological Record, 43*, 725–744.

Saunders, R. R., Chaney, L., & Marquis, J. G. (2005). Equivalence class establishment with two-, three-, and four-choice matching to sample by senior citizens. *The Psychological Record, 55*, 539–559.

Saunders, R. R., Chaney, L., & Marquis, J. G. (2005). Equivalence Class Establishment with Two-, Three-, and Four-Choice Matching to Sample by Senior Citizens. *Psychological Record, 55*, 539–559.

Saunders, R. R., Drake, K. M., & Spradlin, J. E. (1999). Equivalence class establishment, expansion, and modification in preschool children. *Journal of the Experimental Analysis of Behavior, 71*, 195–214. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1284700/pdf/10220930.pdf.

Saunders, R. R., & Green, G. (1999). A discrimination analysis of training-structure effects on stimulus equivalence outcomes. *Journal of the Experimental Analysis of Behavior, 72*, 117–137.

Scandura, J. M. (1971). *An algorithmic approach to mathematics: Concrete behavioral foundations*: Harper & Row.

Schlinger, H. D., & Poling, A. D. (1998). *Introduction to scientific psychology*. New York: Plenum Press.

Schroeder, S. R., & Holland, J. G. (1968). Operant control of eye movements. *Journal of Applied Behavior Analysis, 1*, 161–166. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16795173.

Schroeder, S. R., & Holland, J. G. (1969). Reinforcement of eye movement with concurrent schedules. *Journal of the Experimental Analysis of Behavior, 12*, 897–903. doi: 10.1901/jeab.1969.12-897

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech & Hearing Research, 14*, 5-13. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5550631.

Sidman, M. (1986). Functional analysis of emergent verbal classes. *Analysis and integration of behavioral units*, 213–245.

Sidman, M. (1992a). Equivalence Relations: Some Basic Considerations. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding Verbal Relations*. Reno, NV Context Press.

Sidman, M. (1994). *Equivalence relations and behavior: A research story*: Authors Cooperative.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74*, 127–146. doi: 10.1901/jeab.2000.74-127

Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children. *Journal of the Experimental Analysis of Behavior, 37*, 23–44.

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of behavior, 37*, 5–22. doi: 10.1901/jeab.1982.37-5

Sidman, M., Willson-Morris, M., & Kirk, B. (1986). Matching-to-sample procedures and the development of equivalence relations: The role of naming. *Analysis and intervention in Developmental Disabilities, 6*, 1–19.

Skinner, B. F. (1935). The generic nature of the concepts of stimulus and response. *The Journal of General Psychology, 12*, 40–65. doi: 10.1080/00221309.1935.9920087

Skinner, B. F. (1938). *The behavior of organisms. new york: Appleton-Century-Crofts* (Vol. 221).

Skinner, B. F. (1953). *Science and human behavior*: SimonandSchuster. com.

Skinner, B. F. (1957). *Verbal behavior*. Englewood Cliffs, N.J: Prentice-Hall.

Skinner, B. F. (1961). *Cumulative record (Enl. ed.)*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1968). *The technology of teaching*: BF Skinner Foundation.

Smoke, K. L. (1932). An objective study of concept formation. *Psychological Monographs, 42*, i.

Spencer, T. J., & Chase, P. N. (1996). Speed analyses of stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 65*, 643–659.

Spradlin, J. E., Cotter, V. W., & Baxley, N. (1973). Establishing a conditional discrimination without direct training: a study of transfer with retarded adolescents. *American Journal of Mental Deficiency*.

Spradlin, J. E., Saunders, K. J., & Saunders, R. R. (1992). The stability of equivalence classes. *Understanding verbal relations*, 29–42.

Steingrimsdottir, H. S., & Arntzen, E. (2016). Eye Movements During Conditional Discrimination Training. *The Psychological Record, 66*, 201–212.

Tomanari, G. Y., Balsamo, L. M., Fowler, T. R., Lombard, K. M., Farren, K. M., & Dube, W. V. (2007). Manual and ocular observing behavior in human subjects. *European Journal of Behavior Analysis, 8*, 29–40. Retrieved from http://www.ejoba.org/.

Urcuioli, P. J. (2013). Stimulus control and stimulus class formation.

Watson, J. B. (1930). *Behaviorism* (Rev. ed. ed.). Chicago: University of Chicago Press.

Weiss, A. P. (1925). A theoretical basis of human behavior: by Albert Paul Weiss *Human behavior, A theoretical basis of.*: Columbus, O.: R. G. Adams & co.

Wilkinson, K. M., & McIlvane, W. J. (2001). Methods for studying symbolic behavior and category formation: Contributions of stimulus equivalence research. *Developmental Review, 21*, 355–374.

Yarbus, A. L. (1967). *Eye movements and vision* (Vol. 2): Plenum press New York.

Yerkes, R. M. (1928). The mind of a gorilla: Part III. *Genetic Psychology Monographs, 201*, 256.

Zentall, T. R., Galizio, M., & Critchfield, T. S. (2002). Categorization, concept learning, and behavior analysis: An introduction. *Journal of the experimental analysis of behavior, 78*, 237-248.

Study 1

Fixating, Attending, and Observing: A Behavior Analytic Eye-Movement Analysis

Steffen Hansen and Erik Arntzen

Hansen, S., & Arntzen, E. (2015). Fixating, attending, and observing: a behavior analytic eye-movement analysis. *European Journal of Behavior Analysis, 16*(2), 229–247. doi:10.1080/15021149.2015.1090750

EUROPEAN JOURNAL OF BEHAVIOR ANALYSIS, 2015
http://dx.doi.org/10.1080/15021149.2015.1090750

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.

ARTICLE HISTORY
Received 17 January 2015
Accepted 28 August 2015

KEYWORDS
Stimulus control; attending; observing response; measurements; detection algorithms; eye-fixation; eye-tracking

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

> 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 capability 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 intervening visual discriminative stimuli. A fixation on a comparison stimulus, without the evocation of a discriminative response, suggests that this particular comparison stimulus played the role of S^{Δ} 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

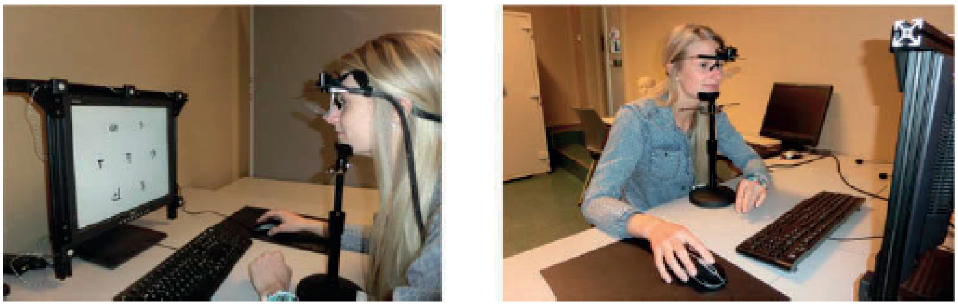


Figure 1. A participant situated in an experimental preparation. Wearing an eye-tracking glass, recording eye-movements of the left eye, the head is supported by and is rested on a non-intrusive chin cup system.

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		A complex integration of light, form, contrast, and color sense.
	Visual acuity	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.

Table 2. Algorithms related to fixation event detection.

Event detection algorithms	Models	Description
Dispersion-based algorithms		Used with low-speed eye-trackers because low velocity data points cluster relatively close to each other.
	Dispersion-Threshold Identification Algorithm (I-DT)	Identifies eye-fixations through the use of contiguous data samples that are located within a predetermined "window-size."
Velocity-based algorithms		Suitable for high-speed eye-trackers because they gather eye-movement data points at a relatively high sampling frequency.
	Velocity-Threshold Identification Algorithm (I-VT)	It separates fixation- and saccade-segments based on their point-to-point velocities.
Probability-based algorithms		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.
	The Hidden Markov Model (I-HMM)	Velocity-based, probabilistic model as it utilizes sequential data segments in its computational protocol.
	Binocular-Individual Threshold Algorithm (BIT)	Velocity-based, probabilistic model as it utilizes sequential data segments in its computational protocol.

Note: The table summarizes the most frequently used algorithms to detect eye-fixation 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.

Eye-fixation

As an event, the structure of eye-fixation is further delineated by (a) the *eye-in-head fixation*, which occurs when the 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), and (b) the *eye-on-stimulus fixation*, which occurs when the 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) (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 fixation events). Event detection algorithms make use of three specific 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 fixation identification 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 fixation events (Salvucci & Goldberg, 2000). The dispersion-threshold identification algorithm (I-DT) identifies eye-fixations through the use of contiguous data samples that are located within a predetermined "window-size." The literature suggests that an observing response to 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).

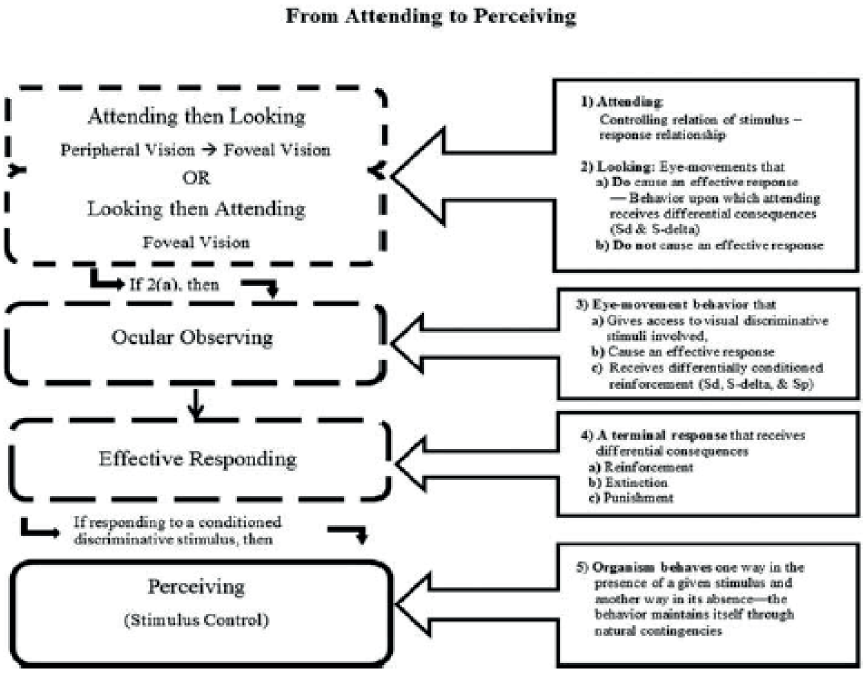


Figure 2. Based on reviewed literature, a flowchart illustrates the controlling relations and behavioral principles that govern an eye's contributing measures to complex human behavior with attending and looking at one end of a continuum and observing and perceiving further along the continuum, respectively.

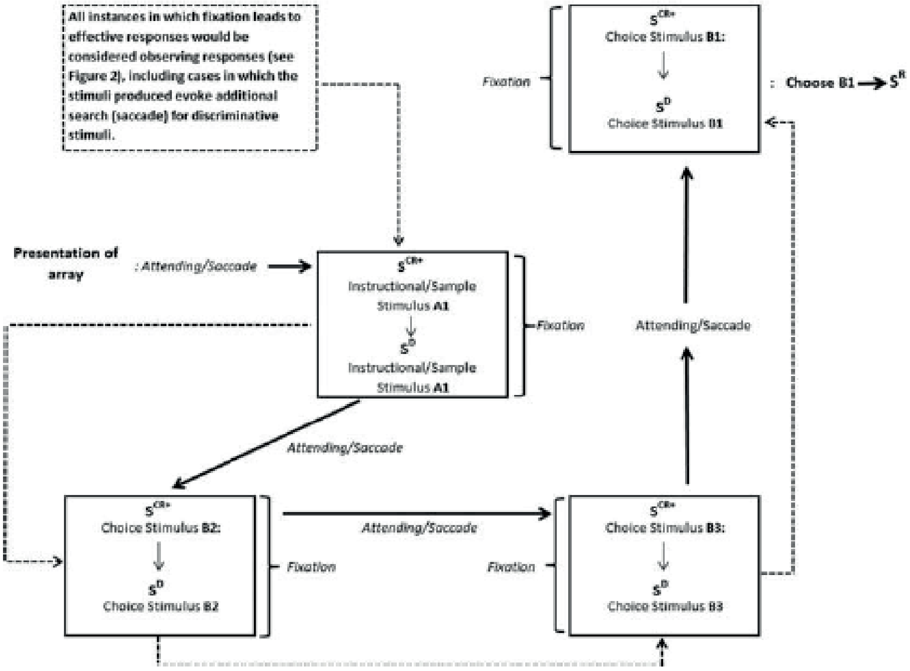


Figure 3. Implicit linkage of the three-term contingencies that are embedded in ocular observing responses during conditional discrimination training—here, a simplified example of matching-to-sample in which the choice of B1 (and not B2 and B3) receives a programmed consequence (S^R) in the presence of instructional (sample) stimulus A1.

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^D) but not in its absence (S^Δ). 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' inabilities 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^Δ stimuli) as the observed responses. Results showed that participants looked at visual S^D and S^Δ 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 Arzuoan (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

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*, 68–72. doi:10.1038/nature03086

Arntzen, E., & Hansen, S. (2013). *Eye-tracking technology as an additional measurement in conditional discrimination procedures.* Paper presented at the Tobii Eye Tracking Conference on Behavioral Research, Boston, MA.

Baum, W. M. (2005). *Understanding behaviorism: Behavior, culture, and evolution*. Malden, MA: Blackwell.

Bennett, A. G., & Rabbetts, R. B. (1989). *Clinical visual optics*. Oxford: Butterworth Heinemann.

Buswell, G. T. (1935). *How people look at pictures: A study of the psychology of perception in art*. Chicago, IL: University of Chicago Press.

Dinsmoor, J. A. (1985). The role of observing and attention in establishing stimulus control. *Journal of the Experimental Analysis of Behavior*, *43*, 365–381. doi:10.1901/jeab.1985.43-365

Dodge, R. (1907). An experimental study of visual fixation. *Psychological Monographs: General and Applied*, *8*, i–95. doi:10.1037/h0093042

Donahoe, J. W., & Palmer, D. C. (2004). *Learning and complex behavior*. Richmond, MA: Ledgetop Corporation.

Dube, W. V., Balsamo, L. M., Fowler, T. R., Dickson, C. A., Lombard, K. M., & Tomanari, G. Y. (2006). Observing behavior topography in delayed matching to multiple samples. *The Psychological Record*, *56*, 233–244. Retrieved from http://opensiuc.lib.siu.edu/tpr/vol56/iss2/6

Dube, W. V., Dickson, C. A., Balsamo, L. M., O'Donnell, K. L., Tomanari, G. Y., Farren, K. M., … McIlvane, W. J. (2010). Observing behavior and atypically restricted stimulus control. *Journal of the Experimental Analysis of Behavior*, *94*, 297–313. doi:10.1901/jeab.2010.94-297

Dube, W. V., Lombard, K. M., Farren, K. M., Flusser, D. S., Balsamo, L. M., & Fowler, T. R. (1999). Eye tracking assessment of stimulus overselectivity in individuals with mental retardation. *Experimental Analysis of Human Behavior Bulletin*, *17*, 8–14. Retrieved from http://www.eahb.org/1/wp-content/uploads/2014/03/eahb-vol-17-Dube-et-al.pdf

Dube, W. V., Lombard, K. M., Farren, K. M., Flusser, D. S., Balsamo, L. M., & Fowler, T. R., & Tomanari, G. (2003). Stimulus overselectivity and observing behavior in individuals with mental retardation. In S. Soraci. Jr. & K. Murata-Soraci (Eds.) *Visual Information Processing* (pp. 109–123). Westport, CT: Praeger Publishers.

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice* (Vol. 373). London: Springer.

Hansen, S., & Arntzen, E. (2013). *A review on the use of eye-tracking technology to study observing behavior during Conditional discrimination training.* Paper presented at the ABAI 39th Annual Conference, Minneapolis, MN.

Henderson, J. M., Weeks, P. A., Jr., & Hollingworth, A. (1999). Eye movements during scene viewing: Effects of semantic consistency. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 210–228. Retrieved from http://psycnet.apa.org/journals/xhp/25/1/210/

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Horsley, M., Eliot, M., Knight, A., & Reilly, R. (2014). *Current trends in eye tracking research.* Switzerland: Springer International Publishing.

Ishizuka, K., Kashiwakura, M., & Oiji, A. (1998). Eye movements in patients with schizophrenia: Visual stimuli, semantic content and psychiatric symptoms. *Acta Psychiatrica Scandinavica*, *97*, 364–373. doi:10.1111/j.1600-0447.1998.tb10016.x

Jenkins, H. M., & Harrison, R. H. (1960). Effect of discrimination training on auditory generalization. *Journal of Experimental Psychology*, *59*, 246–253. doi:10.1037/h0041661

Khurana, A. K. (2008). *Theory and practice of optics and refraction* (2nd ed.). Okhla, New Delhi: Elsevier.

Kirshner, H., & Sidman, M. (1972). Scanning patterns in aphasic patients during matching-to-sample tests. *Neuropsychologia*, *10*, 179–184. doi:10.1016/0028-3932(72)90057-7

Laeng, B., & Teodorescu, D.-S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science*, *26*, 207–231. doi:10.1207/s15516709cog2602_3

Lueck, K. L., Mendez, M. F., & Perryman, K. M. (2000). Eye movement abnormalities during reading in patients with Alzheimer disease. *Cognitive and Behavior Neurology*, *13*, 77–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10780625

Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and Instruction*, *20*, 167–171. doi:10.1016/j.learninstruc.2009.02.012

Nakayama, M., Takahashi, K., & Shimizu, Y. (2002). *The act of task difficulty and eye-movement frequency for the 'Oculo-motor indices'.* Paper presented at the Proceedings of the 2002 Symposium on Eye Tracking Research & Applications, New Orleans, LA.

Nodine, C. F., Locher, P. J., & Krupinski, E. A. (1993). The role of formal art training on perception and aesthetic judgment of art compositions. *Leonardo*, *26*, 219–227. doi:10.2307/1575815

Palmer, D. C. (2010). Behavior under the microscope: Increasing the resolution of our experimental procedures. *The Behavior Analyst*, *33*, 37–45. Retrieved from http://www.ncbi.nlm.nih.gov/

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422. doi:10.1037/0033-2909.124.3.372

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*, 1457–1506. doi:10.1080/17470210902816461

Rayner, K., & Pollatsek, A. W. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice-Hall.

Rudolph, R. L., & Houten, R. (1977). Auditory stimulus control in pigeons: Jenkins and Harrison (1960) revisited. *Journal of the Experimental Analysis of Behavior*, *27*, 327–330. doi:10.1901/jeab.1977.27-327

Salthouse, T. A., & Ellis, C. L. (1980). Determinants of eye-fixation duration. *The American Journal of Psychology*, *93*, 207–234. doi:10.2307/1422228

Salthouse, T. A., Ellis, C. L., Diener, D. C., & Somberg, B. L. (1981). Stimulus processing during eye fixations. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 611–623. doi:10.1037/0096-1523.7.3.611

Salvucci, D. D., & Goldberg, J. H. (2000). *Identifying fixations and saccades in eye-tracking protocols*. Paper presented at the Proceedings of the 2000 Symposium on Eye Tracking Research & Applications, Palm Beach Gardens, FL.

Schlinger, H. D. (2008). *A behavioral perspective on listening and auditory imagining*. Paper presented at the CalABA 26th Annual Western Regional Conference, Garden Grove, CA. Retrieved from https://calaba.org/2008conf/program-search.asp?mode=search

Schlinger, H. D. (2009). Auditory imagining. *European Journal of Behavior Analysis*, *10*, 77–85. Retrieved from http://www.ejoba.org/PDF/2009_1/Schlinger_2009.pdf

Skinner, B. F. (1935). The generic nature of the concepts of stimulus and response. *The Journal of General Psychology*, *12*, 40–65. doi:10.1080/00221309.1935.9920087

Skinner, B. F. (1953). *Science and human behavior*. New York, NY: Simon & Schuster Inc.

Smith, F. (1971). *Understanding reading: A psycholinguistic analysis of reading and learning to read*. New York, NY: Routledge.

Tomanari, G. Y., Balsamo, L. M., Fowler, T. R., Lombard, K. M., Farren, K. M., & Dube, W. V. (2007). Manual and ocular observing behavior in human subjects. *European Journal of Behavior Analysis*, *8*, 29–40. Retrieved from http://www.ejoba.org/

Vakil, E., & Lifshitz, H. (2012). Solving the Raven Progressive Matrices by adults with intellectual disability with/without Down syndrome: Different cognitive patterns as indicated by eye-movements. *Research in Developmental Disabilities*, *33*, 645–654. doi:10.1016/j.ridd.2011.11.009

Vakil, E., Lifshitz, H., Tzuriel, D., Weiss, I., & Arzuoan, Y. (2011). Analogies solving by individuals with and without intellectual disability: Different cognitive patterns as indicated by eye movements. *Research in Developmental Disabilities*, *32*, 846–856. doi:10.1016/j.ridd.2010.08.006

van der Lans, R., Wedel, M., & Pieters, R. (2011). Defining eye-fixation sequences across individuals and tasks: The Binocular-Individual Threshold (BIT) algorithm. *Behavior Research Methods*, *43*, 239–257. doi:10.3758/s13428-010-0031-2

West, J. M., Haake, A. R., Rozanski, E. P., & Karn, K. S. (2006). *eyePatterns: Software for identifying patterns and similarities across fixation sequences*. Paper presented at the Proceedings of the 2006 Symposium on Eye Tracking Research & Applications, San Diego, CA.

Wyckoff, L. B. (1952). The role of observing responses in discrimination learning. Part I. *Psychological Review*, *59*, 431–442. doi:10.1037/h0053932

Yarbus, A. L. (1967). *Eye movements and vision* (Vol. 2). New York, NY: Plenum press.