Journal of Behavioral Decision Making, *J. Behav. Dec. Making*, **29**: 103–115 (2016)
Published online 20 March 2015 in Wiley Online Library (wileyonlinelibrary.com) **DOI:** 10.1002/bdm.1867

Areas of Interest as a Signal Detection Problem in Behavioral Eye-Tracking Research

JACOB L. ORQUIN[1]*, NATHANIEL J. S. ASHBY[2] and ALASDAIR D. F. CLARKE[3]
[1]*Department of Business Administration/MAPP, Aarhus University, Aarhus, Denmark*
[2]*Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA, USA*
[3]*School of Psychology, University of Aberdeen, Aberdeen, UK*

ABSTRACT

Decision researchers frequently analyze attention to individual objects to test hypotheses about underlying cognitive processes. Generally, fixations are assigned to objects using a method known as *area of interest* (AOI). Ideally, an AOI includes all fixations belonging to an object while fixations to other objects are excluded. Unfortunately, due to measurement inaccuracy and insufficient distance between objects, the distributions of fixations to objects may overlap, resulting in a signal detection problem. If the AOI is to include all fixations to an object, it will also likely include fixations belonging to other objects (false positives). In a survey, we find that many researchers report testing multiple AOI sizes when performing analyses, presumably trying to balance the proportion of true and false positive fixations. To test whether AOI size influences the measurement of object attention and conclusions drawn about cognitive processes, we reanalyze four published studies and conduct a fifth tailored to our purpose. We find that in studies in which we expected overlapping fixation distributions, analyses benefited from smaller AOI sizes (0° visual angle margin). In studies where we expected no overlap, analyses benefited from larger AOI sizes (>.5° visual angle margins). We conclude with a guideline for the use of AOIs in behavioral eye-tracking research. Copyright © 2015 John Wiley & Sons, Ltd.

KEY WORDS eye-tracking; areas of interest; signal detection; data fishing; information processing

AREAS OF INTEREST AND AUXILIARY ASSUMPTIONS

Since the seventies, decision research has relied, in part, on measurement of eye-movements to trace cognitive processes. Many shared concepts have emerged and are frequently applied in this line of research, such as the notions of *occurrence* and *adjacency*: that attending to an object is a necessary condition for cognitive processing, and that the order in which objects are attended to reveals the order of processing. Also, the notion of *weight*, that the amount of attention reveals the importance of an object to the decision maker (see for instance Costa-Gomes et al., 2001; Svenson, 1979). By applying these principles, researchers may test hypotheses about decision processes (Russo, 2011), but in doing so must make an auxiliary assumption about the accuracy of object attention measurement. Despite 4 decades of research on eye-movements in decision making, no one has addressed the boundary conditions of this assumption.

The most common approach to determining object attention is by assigning all fixations to an object that fall within a certain distance of the object. This approach is known as classification by *area of interest* (AOIs; also known as regions of interest). There are at least two ways to define an AOI: either using software to physically draw the AOI or through the use of scripts that define and assign fixations or saccades to AOIs based on their coordinates on the visual plane. Both methods are used extensively in the area of behavioral decision making (Schulte-Mecklenbeck et al., 2015). The drawing method is often used for more visually complex stimuli (visual scenes and oddly shaped objects), whereas script definitions are used where more precision is both possible and required (text stimuli and easily defined shapes). Unfortunately, the aspects that make both methods favorable also lead to differences in the data used for analyses and the level of detail in reporting. Importantly, it is often the case that the way in which AOIs are defined is limited by the researchers' programming abilities, as advanced scripts must be constructed for fixation classifications, which are not readily available in analyses packages combined with eye-tracking equipment.

Holmqvist et al. (2011) note that the AOI method has been continuously reinvented by researchers and software developers, leading to a lack of common terminology and methods for defining AOIs. While discipline-specific standards may exist, for instance in reading research (Rayner, 2009), these standards are rarely employed in, or optimally designed for, decision research. In addition, there is no consensus regarding the use and reporting of AOIs in decision making research. This lack of standardization in AOI definition and reporting presents direct problems to the advancement of behavioral decision making, problems we attempt to address in the current manuscript.

Regarding the definition of AOIs, Holmqvist and colleagues (2011) have offered several recommendations. Regarding the size of AOIs, they recommend that AOIs should allow for a buffer space (margin) of 1° to 1.5° of visual angle around the object, an area approximately the size of the fovea. If accuracy is low, then margins can be increased further to ensure inclusion of all fixations belonging to a given object. In situations where the researcher is free to

*Correspondence to: Jacob L. Orquin, Department of Business Administration/MAPP, Aarhus University, Aarhus, Denmark. E-mail: jalo@badm.au.dk

design the stimuli, it is recommended that enough distance around objects be provided so that the researcher can increase AOI margins without risking overlap, minimizing the risk of participants processing more than one stimulus at a time.

In general, the advice from Holmqvist et al. (2011) seems to be to keep AOIs maximal to avoid false negatives. While this advice may be sound in many cases, it ignores situations that would benefit from a more complete signal detection framework. The problem is that by increasing AOI margins to avoid false negatives, one risks attributing fixations that do not belong to an object. The problem is further amplified if objects in the visual scene are close together as fixations landing between objects could belong to either, or both, of the neighboring objects. Related to this is the suggestion that researchers should not change their AOIs after data collection and should instead use standardized AOI definitions for the type of stimuli they are working with. Although there are techniques that decide the AOI size and location based on the observed data, such as clustering algorithms, rendering the AOIs after observing the data could be problematic as it opens up the temptation to engage in data fishing: performing multiple analyses and only reporting those that produce significant *p*-values (Simmons et al., 2011).

The use of AOI margins (Holmqvist et al., 2011) is difficult to ascertain in decision making research as details relating to AOI specifications are frequently omitted (Schulte-Mecklenbeck et al., 2015). In addition, the distance between AOIs is highly variable between studies. In some cases, particularly when annotated photographs have been used as stimuli, targets in the scene can overlap, leaving the experimenter with the problem of deciding how best to define AOIs. To account for this problem of overlap Yun et al. (2013) simply counted the number of fixations falling within each bounding box, thereby allowing fixations to be assigned to more than one AOI. An alternative approach was used by Clarke et al. (2013), who assigned each fixation to the smallest AOI that contained it; unlike most of the studies discussed earlier, they used polygon AOIs that traced the outline of the objects.

To the best of our knowledge, no studies have examined how AOI sizes impact data analysis or the processing assumptions often linked to directed attention. It therefore remains an open question whether the keep-it-maximal approach is suitable for all types of stimuli and analyses used in behavioral decision making research. This paper aims to cast light on the issue of AOI definition. First, we explicate on the problems associated with AOI definition, framing them as common signal detection problems. Then, to situate the issue in what is current practice we report the results of a survey on the use and reporting of AOIs by researchers. We then reanalyze four previously published eye-tracking experiments using different AOI sizes and stimuli types. In addition, we conduct a fifth experiment which directly examines the influence of object distance on the overlap of fixation distributions and cognitive processing. We conclude with a discussion of what the current findings suggest in terms of practical guidelines for AOI definition in decision making research.

AREAS OF INTEREST AS A SIGNAL DETECTION PROBLEM

As mentioned earlier, little, or no, research exists on the use of AOIs in decision making research, and the only advice that has been given is to keep AOIs maximal in order to include all fixations that belong to an object (Holmqvist et al., 2011). However, in an attempt to include all fixations belonging to an object or attribute, one runs the risk of inflating the false positive rate (i.e., including fixations that were not related to the object or attribute). As such, we suggest that it is worthwhile to frame the definition of AOI margins as a signal detection problem. When researchers decide the margin of an AOI they face two problems: If one makes the AOI margin too large, it could include too many false positives, rendering the AOI meaningless as it would cover a semantically heterogeneous area (meaningful AOIs cover semantically homogenous areas). On the other hand, if one makes the AOI margin too small, the statistical power and link to underlying cognitive processing is reduced by excluding relevant observations (see Figure 1). The problem is aggravated by the fact that a researcher cannot precisely know the distribution of fixations to an object. As such, it seems plausible that there must be an AOI margin that optimally balances the false positive and negative rates with statistical power. Although we, de facto, do not know whether this is true, again because the true distribution of fixations is unknown, we suggest that if an optimal AOI margin exists, it will increase the explanatory power of the analysis. In other words, if an AOI is defined so that it better captures attentional allocation and processing, we would expect the predictive power associated with that AOI to increase.

Inspecting eye-movement data, one often finds that fixations are scattered around objects in a cloud as seen in Figure 1. We suggest that at least three different factors may influence the distribution of fixations:

Hardware accuracy

Eye-tracking hardware has an upper bound for how accurate eye-movements can be recorded; the actual accuracy may

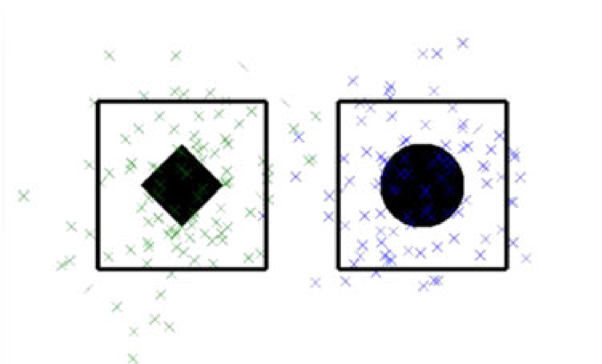

Figure 1. Hypothetical example of the signal-detection problem: The areas of interest (AOIs) (boxes) do not include all fixations to each object (false negatives) because extending the AOI margins would include fixations belonging to the other object (false positives)

DOI: 10.1002/bdm

vary as a function of the calibration, which we note is generally less accurate on the y axis than the x axis and in the corners of the screen, leading to an even larger distribution of fixations. Hardware accuracy is, to a large extent, under the control of the experimenter who is free, barring funding constraints, to acquire hardware with better accuracy, define calibration as being stringent (liberal), and by re-calibrating participants at regular intervals.

Saccadic accuracy

Saccades (eye-movements) often undershoot their target and are followed up by a corrective saccade that positions the eyes more directly on the target (Palmer, 1999). Accurate saccades take more time to program, and the eye-movement system may trade-off accuracy for speed (Kowler, 2011). Saccadic accuracy may be hard for the experimenter to control, but certain conditions, such as time pressure, could lead to a decrease in saccadic accuracy. Furthermore, the eyes may drift during fixations producing a slow involuntary movement away from the target object often paired with a corrective, reorienting saccade (Steinman et al., 1973).

Center of gravity averaging

Human observers often fixate in the middle of a group of objects, an effect that has been labeled the center of gravity averaging response (Zelinsky, 2008). From the central viewing position, more objects may be processed simultaneously, and the observer may use one or more center of gravity saccades to localize a target. Center of gravity saccades are more prevalent for observers with larger perceptual span (Reingold et al., 2001). The experimenter may have some influence over center of gravity saccades, for instance, by controlling the distance and clustering of objects, the discriminability of objects in parafovea through size and contrast (Anstis, 1974; Melmoth & Rovamo, 2003), crowding (Pelli et al., 2007), and domain-specific expertise, which enhances the perceptual span (Reingold & Sheridan, 2011).

In this paper, we analyze several experiments using different AOI sizes in order to learn more about the putative distribution of fixations. The endeavor is complicated by the fact that we cannot directly observe the distribution of fixations, nor can we know for sure how the distributions are influenced by experimental factors, or how they influence our results given a particular AOI margin. That being said, we have several general expectations. Specifically, we expect that lower eye-tracking accuracy, smaller distances between objects, larger and more salient stimuli, and familiarity with experimental stimuli will increase the overlap of fixation distributions. We also expect that experiments with a higher overlap of fixation distributions will benefit from smaller AOI margins, whereas experiments with little or no overlap of fixation distributions will benefit from larger, more inclusive, AOI margins.

Table 1. Responses to questions about the use of areas of interest (AOI).

Question	Yes	No	Do not know
AOI same size as stimuli?	47%	53%	—
AOI larger than stimuli?	61%	39%	—
AOI smaller than stimuli?	14%	84%	2%
Allow AOIs to overlap?	21%	75%	4%
Account for eye-tracking accuracy?	44%	53%	4%
Account for perceptual span?	25%	68%	7%
Employ fuzzy AOIs?	11%	84%	5%
Analyze multiple AOI sizes?	35%	65%	—

CURRENT PRACTICE

Before commencing the analysis of the experiments, we considered it useful to survey the current practice among eye-tracking researchers with regard to the use of AOIs. No reviews or summaries exist on current practice, and because AOIs are frequently underspecified in the literature, it is difficult to ascertain how AOIs are currently used in eye-tracking research, particularly in more junior areas such as behavioral decision making. Hence, a survey should not only reveal whether researchers are following the advice of Holmqvist et al. (2011) to keep AOIs maximal but also allow us to compare current practice with optimal practice through the reanalysis of published data. To this end, we launched an online survey containing several questions on the use of AOIs. The survey was distributed through mailing lists belonging to research communities such as judgment and decision making, economics, psychology, marketing, and neuroscience.[1]

Seventy-one individuals responded to our survey with 80% of individuals indicating that they had familiarity with publishing eye-tracking research; that is, they had at least one manuscript under review or published. Notably, 50% of individuals reported that they had two or more manuscripts either published or submitted for publication, with 27% reporting five or more. Because our aim is to get a general feel for the types of analyses being used by active researchers (researchers seeking to publish their results), we include only individuals who had submitted at least one manuscript for publication resulting in 57 responses being retained. Respondents came from a wide range of research backgrounds with 3% identifying as psychologists, 22% as economists/marketers, 14% as judgment and decision making researchers, 46% as neuroscientists, 2% as commercial researchers, and 13% indicating some other area of research.

There was quite a deal of variation in how our respondents defined AOIs in their studies. Looking at Table 1, we see that our respondents reported using AOIs both the same size as, smaller than, and larger than the stimuli of interest; 21% of respondents reported using overlapping AOIs and 11% employed fuzzy AOIs (i.e., weighting fixations to

[1]The survey was advertised on the following mailing lists/societies: Society for Judgment and Decision Making, Society for the Advancement of Behavioral Economics, European Association for Decision Making, Society for Consumer Psychology, Eye Movement Mailing List, and Vision Sciences Society.

DOI: 10.1002/bdm

reflect how close they are to different AOIs). Overall then, in terms of our respondents, it does not seem that there is a standard AOI definition being employed. In fact, we find that over half of our respondents did not report taking into account the accuracy of their eye-tracking equipment when forming their AOIs.

What then do we learn from the survey? First, there appears to be no standard practice being followed by our small sample of respondents. Although 61% answered that they used AOIs larger than the stimulus as recommended by Holmqvist et al. (2011), roughly half of respondents reported using AOIs the same size as the stimulus, and 14% indicated they had used AOIs smaller than the stimulus. Second, the survey suggests that the use of AOIs may be somewhat unreflective of the auxiliary assumptions behind comparisons of object attention: The majority of researchers do not consider the accuracy of their eye-tracking hardware or perceptual span, both of which could lead to inflated false positive and negative rates. Finally, roughly a third of responders analyzed more than one AOI size, which could be an indication of data fishing depending on what the researcher chose to report (Simonsohn et al., 2014; Simmons et al., 2011). Importantly, our respondents represent only a fraction of those researchers currently engaged in eye-tracking research, and there is no way for us to measure their particular impact on the field (i.e., are they regarded as experts). We therefore urge caution in making generalizations about all eye-tracking researchers based solely on these findings. Nevertheless, the current survey suggests that many eye-tracking researchers would benefit from an investigation into the definition and use of AOIs.

DO AREA OF INTEREST SIZES INFLUENCE OUR RESULTS?

The previous section suggests that current practice regarding the use of AOIs may be suboptimal in several ways. It is clear that practices such as analyzing different AOI sizes for the purpose of producing significant results would be best avoided completely, while other practices such as ignoring the perceptual span, or the accuracy of the eye-tracking equipment, may constitute lesser problems. In this section we reanalyze four different experiments using varying AOI sizes. This allows us to examine whether AOI size influences the significance and predictive power of the results. The experiments are based on different tasks and use different classes of stimuli that range from being visually simple (matrices with numbers or text) to highly complex (photographs of scenes). Additionally, we conduct a fifth eye-tracking experiment in which we systematically manipulate the distance between stimuli; no previous studies that we reviewed were designed to examine the effects of stimulus distance.

All studies were analyzed using at least three, and in some cases five, different AOI margins. The number of margins used in each analysis depends mainly on the characteristics of the stimuli. Whereas some experiments may use stimuli with large distances between objects, allowing AOI margins of up to 2° of visual angle, other studies may, for instance, only allow for margins of 1° of visual angle before overlap occurs. For consistency, each of the four studies is reanalyzed using the same, or a similar, model as in the original paper in which the data were first presented.

Study 1

In Study 1, we examine the effect of AOI size in the valuation of risky prospects. In their third study, Ashby et al. (2012) had participants who provide valuations of 80 risky prospects (gambles) that consisted of two non-negative outcomes (e.g., 3% chance of 9.70 Euro or 97% chance of 0.30 Euro) as either buyers or sellers. Outcomes and their respective probabilities were presented as text on the left and right hand sides of the screen. Eye-movements were recorded using the Eyegaze binocular system (LC Technologies, VA, USA), which provided for binocular sampling at a rate of 120 Hz and provided an accuracy of approximately 0.45° of visual angle. Fixations were defined using a script developed by the authors. The visual layout of their study can be characterized as simple with little possibility of overlapping fixation distributions.

In terms of the hypothesis tested, the authors predicted, and found, that the gaze duration to an outcome was predicted both by the probability of that outcome occurring, as well as whether the valuation was made as a buyer or as seller. Specifically, higher probability outcomes were attended to more frequently, and buyers focused to a greater extent on low outcomes relative to sellers. In addition, the authors hypothesized, and found that the proportion of time spent attending to the low outcome was predictive of the valuation given, with more attention to low outcomes predicting lower valuations.

Area of interest definition

Ashby et al. (2012) presented stimuli with a width of approximately 1.14° and a height of approximately 1.42°. Their reported AOIs, however, were 2.54° × 2.83° (100 × 100 pixel squares centered on each outcome and probability combination). Thus, their AOIs provided approximately .7° margin in terms of width and height. In terms of the accuracy of their eye-tracking equipment, their AOIs allowed for margins about one and a half times the best possible accuracy of the eye tracker. In the current study, we reanalyze the data of Ashby et al. (2012) and compare the published results with those obtained using AOIs with 0°, .5°, 1°, and 1.5° of visual angle margin around the stimuli.

Reanalysis

Examining the subjective valuation data published by Ashby et al. (2012), we find little difference in the predictive power of both probabilities and perspective on the allocation of visual attention when employing differently sized AOIs (see Table 2). AOIs of 0° and .5° of visual angel are found to lead to marginal results for the effect of perspective, while the effect of probability remains significant. For the original AOIs defined by the authors, as well as the 1° and 1.5° of visual

DOI: 10.1002/bdm

Table 2. Effect by area of interest (AOI) margin for the reanalysis of Ashby et al. (2012) by AOI margin. Original is the AOI definition used by Ashby et al.

AOI margin	Analysis 1			Analysis 2	
	Probability	Perspective	R^2	Valuations	R^2
0°	$b = .291$; $p < .01$	$b = .059$; $p = .092$	.131	$b = -3.100$; $p < .01$	.444
.5°	$b = .288$; $p < .01$	$b = .066$; $p = .053$	.154	$b = -3.748$; $p < .01$	.454
Original	$b = .286$; $p < .01$	$b = .067$; $p = .047$	.155	$b = -3.729$; $p < .01$	.453
1°	$b = .286$; $p < .01$	$b = .068$; $p = .047$	.155	$b = -3.729$; $p < .01$	.453
1.5°	$b = .286$; $p < .01$	$b = .068$; $p = .047$	.155	$b = -3.729$; $p < .01$	.453

angle AOIs, we find no difference in the amount of variance explained or the levels of significance.

In addition, the ability of attention to predict valuations was relatively unaffected by the size of the AOIs with R^2 being highly similar in each case. Thus, at least in the current analysis, the size of the AOIs employed appears to have little effect on the results. Note, however, that AOIs, which allow for a greater border than the absolute accuracy of the eye tracker, provided a slight improvement in predictive power.

Discussion

In Study 1 we analyzed data from Ashby et al. (2012) using five different AOI margins. The results did not differ substantially, but analyses revealed poorer predictive power when the AOI margin was smaller than or just encompassed the best accuracy of the eye tracker (.45° of visual angle). Considering the relatively simple visual design and the large distance between the objects in the display, this could suggest that the distribution of fixations over the two objects in the visual scene did not overlap. It seems then, that in this particular experiment, the researcher is well advised to keep AOI margins maximal as proposed by Holmqvist et al. (2011).

Study 2

In Studies 2a and 2b, we examined the effect of AOI size on attentional allocation in choices between risky prospects (gambles) reported in Fiedler and Glöckner (2012). In these studies, participants made 50 choices between two gambles, both of which contained two possible outcomes. Participants were presented with probabilities and pay-offs for each gamble which were displayed as text and placed in counter balanced order around the edge of an elliptic circle with equal distances between the numbers.

Each trial thus contained eight numbers, and as in the study by Ashby et al. (2012), the display can be characterized as being visually simple. Fiedler and Glöckner (2012) recorded eye-movements using the same eye-tracking system and general fixation detection scripts as Ashby et al. (2012) and tested similar predictions. Namely, Fiedler and Glöckner (2012) predicted attentional allocation by the stated probabilities, outcomes, and their interaction and found support for each, influencing the direction of the gaze.

Area of interest definition

Fiedler and Glöckner (2012) reported AOIs of $2.54° \times 2.83°$ of visual angle (100×100 pixel square centered on each outcome/probability), which were used to fit stimuli of $2.13° \times 1.11°$ of visual angle. Thus, their AOI definitions resulted in an approximately .85° of visual angle margin for height and a .21° visual angle margin for width. Thus, the height of their AOIs was approximately double the best possible error of the eye tracker (.45°), while the width of their AOIs was smaller than the best possible accuracy of the eye tracker. In the current study, we reanalyze the data published by Fiedler and Glöckner (2012) using AOIs, which provide for a margin around each option of 0°, .5°, 1°, and 1.5° of visual angle.

Reanalysis

Examining the risky choice data published by Fiedler and Glöckner (2012), we find very little difference in the predictive power of both probabilities and outcomes on attention when employing various AOI margins (see Table 3). However, as in Study 1, AOIs, which are larger than both the stimuli and the absolute accuracy of the eye tracker, provided a slight improvement in predictive power, although only in Study 2a.

Discussion

Study 2a revealed a small effect of AOI margin on model results. There seems to be a tendency that larger AOI margins lead to more explanatory power for simple text displays, replicating the results of Study 1. Study 2b revealed no effect of AOI margin on predictive power, a surprising finding considering the similarity between the two studies. It is, however, worth noting that the model's estimated parameters increased systematically with larger AOI margins both in Study 2a and 2b.

Study 3

In Study 3, we analyze data from a multi-attribute choice experiment by Orquin, Bagger, and Mueller Loose (2013) on the effect of learning on information reduction. Information reduction refers to the ability to fixate important and ignore unimportant information. The study used a four option multi-attribute choice task with 48 trials in which decision makers made choices between yogurt products according to the attributes: brand, flavor, fat percentage, organic label,

DOI: 10.1002/bdm

Table 3. Effect by areas of interest (AOI) margin for the reanalysis of Fiedler and Glöckner (2012). Study 2a top, Study 2b bottom. Original is the AOI definition used by Fiedler and Glöckner.

AOI margin	Probability	R^2	Outcome	R^2	Probability X Outcome	R^2
0°	$b = 1.04$; $p < .001$	.850	$b = .05$; $p < .001$	.850	$b = .06$; $p = .023$	.520
.5°	$b = 1.48$; $p < .001$	.850	$b = .07$; $p < .001$	.850	$b = .07$; $p = .036$	.548
Original	$b = 1.53$; $p < .001$	.850	$b = .07$; $p < .001$	.850	$b = .07$; $p = .043$	.442
1°	$b = 1.67$; $p < .001$	.850	$b = .08$; $p < .001$	.850	$b = .08$; $p = .025$	.410
1.5°	$b = 1.62$; $p < .001$	.850	$b = .08$; $p < .001$	.850	$b = .08$; $p = .024$	.500
0°	$b = 1.14$; $p < .001$	.648	$b = .07$; $p < .001$	.648	$b = .09$; $p < .001$	.648
.5°	$b = 1.63$; $p < .001$	.648	$b = .09$; $p < .001$	.648	$b = .12$; $p < .001$	.648
Original	$b = 1.71$; $p < .001$	.648	$b = .09$; $p < .001$	.648	$b = .13$; $p < .001$	.648
1°	$b = 1.83$; $p < .001$	.648	$b = .09$; $p < .001$	.648	$b = .13$; $p < .001$	.648
1.5°	$b = 1.93$; $p < .001$	.648	$b = .09$; $p < .001$	.648	$b = .14$; $p < .001$	.648

health label, and price. The experiment contained three between-subjects conditions each of which displayed the information in different ways. In the verbal matrix condition, decision makers were presented with information in a verbal form inserted in a 4×6 matrix. In the visual matrix condition, all information in the matrix, except numeric information (e.g., price), was replaced by images. In the product representation condition, the information was presented in a more naturalistic form resembling actual consumer products. Eye-movements were recorded using a Tobii 2150 (Stockholm, Sweden), which provided for binocular sampling at 60 Hz and an accuracy of approximately 0.7°. Fixations were determined using the Tobii fixation filter.

Relative to Study 1 and 2, this experiment used a less accurate eye tracker and had shorter distances between objects. In addition, the product representation condition used large and salient objects that may have been identifiable outside the fovea. Given these factors, we expect some overlap of the fixation distributions and hence better explanatory power with smaller AOI margins.

Area of interest definition

The original data in the work of Orquin et al. (2013) were extracted using Tobii Studio, a software package that allows the researcher to physically draw AOIs around the stimuli. While this approach allows for a large degree of flexibility regarding the size and shape of AOIs, it provides little flexibility when it comes to changing AOI margins. We therefore extracted the raw data from the eye tracker and computed new AOIs using a script. The original AOIs were rectangular with approximately 1° visual angle margins around each stimulus. The reanalysis used rectangular AOIs with 0°, 0.5°, and 1° of visual angle margin.

Reanalysis

For the analysis, we computed the correlation between fixation likelihood and attribute importance for each AOI within-subjects, providing a measure of the amount of information reduction on a trial-by-trial basis. The importance of the attributes was estimated individually using a random utility model (Louviere et al., 2000) based on each participant's choices in the 48 trials. We then regressed the trial order on information reduction for each of the three conditions. The analysis was carried out for each AOI margin separately and the results are shown in Table 4.

Discussion

Unlike Studies 1 and 2 where stimuli were smaller and with larger distances between them, we believe that Study 3 may have led to overlapping fixation distributions due to lower hardware accuracy, shorter distances between objects, and large and salient stimuli. The analysis revealed a strong positive influence of smaller AOI sizes on the effect of trial order on information reduction, supporting the prediction that smaller AOIs serve to reduce the number of false positive fixations.

Study 4

In Study 4, we examine an extreme case of AOIs using annotated images. In this study, photographs are used as stimuli with the aim of understanding how different visual and semantic factors influence different behavioral tasks. This type of study is becoming increasingly popular as vision scientists try to move to more naturalistic stimuli and tasks (Einhäuser et al., 2008; Nuthmann & Henderson, 2010; Clarke et al., 2013; Yun et al., 2013). In image annotation studies, the AOIs are the objects in a scene. Sometimes there are a small number of target objects, and often an effort is made to annotate all objects present in the scene. There are a number of problems related to the annotation of cluttered scenes, such as dealing with partially occluded objects, and how to assign fixations to overlapping AOIs.

In Study 4, we explore some of these issues in the reanalysis of Clarke et al. (2013). In their study, participants viewed images for 5 seconds. The image was then removed from the display and participants were asked to name as

Table 4. Effect of learning on information reduction (matching fixation likelihood to attributes given individual preferences for that attribute).

AOI margin	Verbal		Visual		Product	
	R^2	p	R^2	p	R^2	p
0°	.19	<.001	.56	<.001	.28	<.001
.5°	.15	.01	.55	<.001	.18	<.001
1°	.04	.15	.52	<.001	.09	.04

DOI: 10.1002/bdm

many objects from the scene as they could remember. The likelihood of a given object being named was modeled by a mixed-effect model with features representing the object's position, salience, linguistic factors, and the amount of overt attention directed towards the object. A collection of 100 images were used as stimuli. Images represented a wide variety of scenes, with each scene containing a median of 26 objects, and representing between 7 and 33 unique object categories. Eye-movements were recorded using an Eyelink II (Ottawa, Ontario, Canada), which provided for monocular sampling at 500 Hz and an accuracy of approximately 0.5° of visual angle. Fixations were determined using the Eyelink II event detection software.

Reanalysis

We investigate how changing the AOI specification affects the conclusions drawn from the mixed-effect model in Section 3.3 of the work of Clarke et al. (2013). In the original model, there were four independent variables: *pos* (representing the position of an object, defined as the Euclidean distance from the center of the image to the center of the object's AOI); *sal* (the visual salience of the object); *lin* (linguistic properties of the associated noun, such as its lexical frequency); and *att* (attention). We chose not to reuse the original attentional scores based on attentional landscapes as these can be thought of as a form of fuzzy AOI and are outside the scope of this paper. Instead, we uniquely assign each fixation to the smallest AOI that contains it to avoid situations in which a fixation is counted multiple times.

We examine the effect of AOI size by applying margins to them. This will cause the *att*, *sal*, and *pos* features to vary as they are defined over the pixels contained within an object's AOI. We replicate the results from Clarke et al. (2013) using 0° visual angle margins with all significant factors having the same sign as in the original analysis.

Next, we explore the effect of increasing AOI margins on model results using the aforementioned same model specification. We find that using AOI margins of approximately .5° of visual angle improve the fit of the model, while 0° of visual angle margins provide the best performance in terms of accuracy and F1[2] (see Table 5).

Discussion

Study 4 differs from the other studies we have reanalyzed as the stimuli were not manipulated by the experimenter. Using photographs and other highly naturalistic types of stimuli represents new problems for the experimenter. First, the objects in the visual scene are often overlapping. Second, there is little distance between objects to allow for additional increases in AOI margin. Initial analysis of this dataset (not presented) suggested that simply counting the number of fixations falling within each AOI, as opposed to uniquely assigning each fixation to the smallest containing AOI, can lead to different patterns of positive and negative effects. In particular, the *att*pos* interaction switches from a significant negative effect to a marginally significant positive effect. While this issue lies beyond the scope of the current study, it suggests that more research is required to determine the most appropriate method for constructing AOIs in complex environments.

We explored the effect of different AOI margins on model fit. Considering that the experiment used naturalistic stimuli where objects vary in size and salience, and where distances between objects are often small or non-existent, it seems plausible that the distribution of fixations to each object will overlap to some degree. The best model fit was obtained with an AOI margin of .5° of visual angle, while 0° of visual angle margins gave the best accuracy, as predicted given an overlap in fixation distributions.

Table 5. Log likelihood of lmer model fits, accuracy, and F1-scores for different area of interest (AOI) margins; an increase of 0.5% in accuracy corresponds to correctly predicting a further 170 named objects.

AOI margin	Log likelihood	Accuracy	F1
0°	−17 283	75.84%	.252
.5°	−17 159	75.78%	.251
1.0°	−17 314	75.54%	.246
1.5°	−17 583	75.22%	.238

Study 5

In the previous studies we find that the effect of the AOI margin on predictive power differs depending on the specifics of the particular experiment. We hypothesized that factors that would lead to a larger spread in the distribution of fixations would, in combination with shorter distances between stimuli, lead to an overlap of fixation distributions. We also hypothesized that given an overlap of fixation distributions, smaller AOI margins will yield better model results (predictive power). Our findings thus far provide some support for these hypotheses. Specifically, we find that in studies in which we expected an overlap of fixation distributions (Studies 3 and 4), predictive power increased with smaller AOI margins, while experiments where we did not expect much overlap, predictive power increased with larger AOI margins (Studies 1 and 2).

Although Studies 1 through 4 confirmed our expectations, none of them provide the opportunity to directly assess the putative importance of object distance in generating overlapping fixation distributions. To remedy this, we conducted a fifth experiment manipulating the distance between choice options in a three-option choice task. Given that shorter object distances should lead to an overlap of fixation distributions it should be possible to detect this in our analyses. First, we predict that shorter distances between objects should lead to increased overlap in fixation distributions, which should diminish the gaze bias, that is, looking more frequently or longer at the chosen option (Orquin & Mueller Loose, 2013; Shimojo et al., 2003).

[2]F1 scores are the measure of a test's accuracy and can be interpreted as a weighted average of the model's precision and recall. Values range from 0 (worst) to 1 (best).

DOI: 10.1002/bdm

Second, given that shorter object distances should result in overlapping fixation distributions, we hypothesize that this will influence models that are sensitive to the processing of choice information. More specifically, it has been demonstrated that choice processes can be described as accumulation processes in which a decision maker accumulates evidence in favor of a choice option during fixations to that option (Krajbich et al., 2011). Thus, the more, and longer, the fixations a decision maker has to an option, the higher the likelihood that they will choose that option. However, if the fixation distributions overlap, we should expect more false negatives. This should lead to options, which are misclassified as not being the current focus of attention, having evidence about them accumulated as if they actually are the focus of attention. Put simply, we predict that when options are close to one another evidence will be accumulated for all options regardless of how a fixation is classified, but when options are far away from one another, evidence will only be accumulated for the option classified as being fixated on.

To test these predictions, we formulate and employ the following simplified discounted evidence accumulation model for choice prediction:

$$\text{If fixation at time } t \text{ is to option } i: \quad V_{i(t)} = V_{i(t-1)} + v_i{*}e_t \tag{1a}$$

$$\text{For all unattended options } j \neq i: \quad V_{j(t)} = V_{j(t-1)} + \theta v_j{*}e_t \tag{1b}$$

The value of a given option (V_i) at the start of the deliberation process (when time $t=0$) is zero. The V_i when an option is being attended to at each subsequent t is calculated by the subjective value (v_i) multiplied by the gaze duration in seconds (e_t) and the previous V_i (Equation 1a). When an option is not the focus of gaze, Equation 1b is used and evidence ($v_i{*}e_t$) for that option is still accumulated, but at a discounted rate equal to . Where can take any value between zero and one, with a value of zero indicating no accumulation of evidence for non-attended to options and a of one indicating constant accumulation of evidence for all options regardless of where attention is directed. The model predicts that the option with a higher accumulated value (V_i) is selected.

METHODS

Participants

Fifty participants ($M_{age}=21.12$, 58% women) with normal to corrected normal vision took part in Study 5 and an unrelated experiment. Participants were paid £5.00 for their participation and the completion time for both experiments was on average 45 minutes.

Materials and procedure

The stimuli used in the experiment consisted of 24 different color photographs of bags of chips (e.g., *Lays Lightly Salted* and *Doritos Cool Ranch*). Upon arriving in the lab, participants were seated (70 cm from the monitor) and read instructions from a 21-inch computer monitor (1680 × 1050 resolution) informing them that they would be providing appetitiveness ratings for snack foods. They were then presented with each of the 24 photos, one at a time in random order. Appetitiveness ratings were indicated by using a mouse to move a slider along a horizontal line (1000 pixels in length) with its ends labeled '*Not at all Appetitive*' and '*Very Appetitive*'. After providing ratings for each item, participants engaged in the unrelated experiment which lasted on average 30 minutes.

Participants were then calibrated, using the default maximum of 9 points, on the eye tracker. Eye-movements were recorded using a SensoMotoric (SMI, Teltow, Germany) RED 250 Hz with a .40° of visual angle accuracy; calibration was accepted if less than 1° of visual angle error (both on the x and y axis) was achieved. Participants then read instructions informing them that they would be making choices between snack food items. The items presented were the same pictures of bags of chips that they had previously rated. Three items (3.43° × 3.48° of visual angle in size) were presented centered on the vertices of one of two equilateral triangles (not visible to participants) centered on the screen. This insured that the center of each item were equidistant from one another. One triangle had side lengths of 5.89° of visual angle while the other had side lengths of 17.19° of visual angle; we term these the *close* and *far* choice arrays, respectively. Items were randomly assigned to locations (vertices) and trials; each item was only shown in one choice set (Figure 2). Participants selected items by rolling the mouse wheel up (down), which moved a small blue outline around the items from one item to the next in clockwise (counter-clockwise) fashion; participants clicked either mouse button once the item they wanted was outlined. In total, eight choices were made, four from each type of array (close or far).

Figure 2. Overlaid screen shot of close and far choice trials in Study 5 with grey outlines representing the approximate areas of interest of 1° of visual angle

DOI: 10.1002/bdm

We defined our AOIs to provide a 1° of visual angle margin around each item reflecting the minimum accuracy achieved during calibration. Fixations were defined as being the total duration of consecutive recordings of gaze with coordinates falling within one of our three AOIs, or on undefined (white) space, and were classified using a script developed by the authors

RESULTS

Gaze bias

Before examining whether the close array diminished the gaze bias, we checked whether the two conditions were comparable in terms of number of fixations. We performed a Poisson regression predicting the number of fixations to stimuli by the spacing of the stimuli. In this and the following analysis, we cluster over participants to control for repeated measurements (Rogers, 1993). In doing so, we find there to be no significant difference in the number of fixations made in the close ($M = 13.67$; $SE = .89$) and far ($M = 15.19$; $SE = 1.74$) arrays, $p = .45$, indicating that the two conditions are comparable in terms of the amount of information acquisition.

Next, to test whether decreasing the distance between the options diminished the gaze bias, we regressed the proportion of fixations to the chosen option on the spacing of the stimuli controlling for options locations. In doing so, we find that the proportion of fixations to the chosen option increases in the far ($M = .45$; $SE = .01$) compared with the close ($M = .39$; $SE = .01$) choice arrays, $b = .05$, $t(49) = 3.24$, $p < .01$. The effect of location was also found to be significant with the option located at a higher point on the screen having a greater proportion of fixations directed at it ($M_{UpperCenter} = .50$) than when the selected option was one of the options presented lower in the display ($M_{LowerLeft} = .38$; $M_{LowerRight} = .35$), both ps $< .001$. Thus, as is often reported, there is a bias in fixations for options that appears in the upper portion of a display. Figure 3 shows a scatter plot of the fixation distributions in the close and far arrays. While the far array results in three separate clusters corresponding to each option the close array results in one large cluster over all three options. Considering the scatter plot and the diminished gaze bias, we suggest that the close array cluster is likely best described as consisting of overlapping fixation distributions.

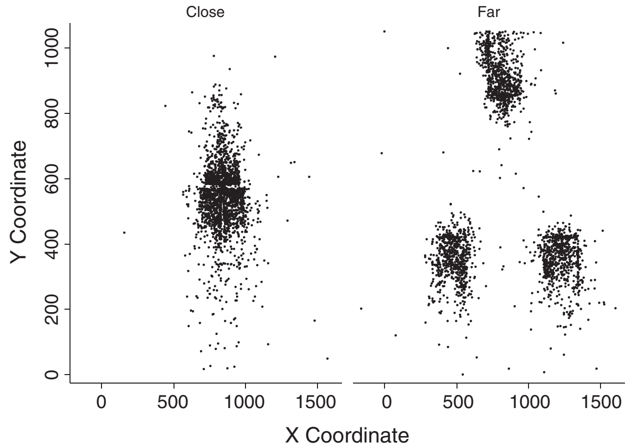

Figure 3. Scatter plot of all fixations greater than 50 ms in the close and far arrays with *X* and *Y* coordinates in pixels

Cognitive processing

We first removed all trials in which an individual had chosen an item they had assigned an appetiveness rating of 0 as such choices could not be fit. We then fit θ on the level of participant by conducting a complete grid search over the parameters space from 0 to 1 in .01 increments and maximizing the log likelihood following the Luce Choice Rule (Luce, 1956) separately for both the close and far arrays. We find that the θ ($M = .09$; $Mdn. = 0$; Range = 0–1) for the close arrays is slightly lower than the θ ($M = .10$; $Mdn = 0$; Range = 0–1) estimated for the far arrays. To see if this difference amounted to a significant reduction, we performed a paired-samples *t*-test and fail to find a significant difference between the two estimated θ's, $t(49) = -.04$, $p = .97$. Thus, counter to what one would predict if overlapping fixation distributions played a role in the evidence accumulation process (i.e., larger θ for close arrays), we fail to find any effect of spacing between options on the rate of accumulation for options not classified as being the focus of gaze.

While not central to the investigation reported here, it is worth noting that our discounted evidence accumulation model did yield more accurate predictions than the appetitiveness ratings alone. Specifically, if we look at the percent of correct predictions using only the appetiveness ratings (option with highest value predicted), we see that, in both the close (69%) and far (67%) arrays, less correct predictions then were made by the discounted evidence accumulation model, which made about 10% more correct predictions in both the close (78%) and far (79%) arrays. Thus, the current study provides some support for the contention that attentional allocation is important in the construction of preference and that gaze can be used to make more accurate choice predictions.

Discussion

In Study 5, we aimed to test whether the putative effect of object distance on the overlap of fixation distributions could be manipulated experimentally. We hypothesized that shorter distances would lead to overlap of fixation distributions and that this overlap would diminish the gaze bias towards the chosen option. As hypothesized, the distance between objects did diminish the gaze bias providing support for the interpretation of AOI margins as a signal detection problem. We also hypothesized that overlap due to shorter object distances would influence the estimates of models designed to fit cognitive processes based on fixation data. In order to test this, we fitted a discounted evidence accumulation model as this model predicts that the likelihood of choosing an option increases when the option is being fixated. Consequently, such a model should suffer when the fixations assigned to each option contain a higher proportion of false positives.

DOI: 10.1002/bdm

However, we find no significant decrease in the estimated discounted accumulation parameter (θ) going from close to far arrays, and that the predictive accuracy of the model remains high in both array types. Therefore, it appears that the spacing of stimuli has little impact on the model estimates, which suggests that for this particular process model overlapping fixation distributions may not be a problem. We note however, that it is possible that more complex models that take into account factors such as object eccentricity would be able to detect other cognitive processes which are affected by overlapping fixation distributions. In addition, it is possible that the stimuli used and a participant's familiarity with the items may have affected processing. As such, we suggest future research should more thoroughly explore such relations.

GENERAL DISCUSSION

Eye-movements are frequently used to trace cognitive processes in decision making. A common approach is to test different theories about the underlying processes by comparing attention to individual objects or option attributes. Such analyses assume that attention is measured accurately across objects, and a typical approach involves assigning fixations to objects using AOIs. However, the use of AOIs has, to the best of our knowledge, not yet been thoroughly examined and no empirical studies of which we are aware have attempted to understand how different AOI definitions may influence such analyses.

The seminal text on AOIs prescribes that researchers use an AOI margin of 1–1.5° of visual angle, and that the margin is increased if the distribution of fixations exceeds the 1.5° of visual angle margin (Holmqvist et al., 2011). Keeping AOIs maximal may be a good idea under some circumstances, but we suggest based on the current results that some situations may require a different approach. We propose that the AOI margin is best understood as a signal detection problem. Specifically, we suggest that the AOI boundaries can be seen as the thresholds for determining the rate of true and false positive fixations that are assigned to any given object, with larger boundaries increasing true but also false positives. We suggest that the distribution of fixations to an object could be influenced by factors such as the accuracy of the eye tracker, the accuracy of saccades, and the frequency of center of gravity saccades. The latter may depend on the size, salience, and distance of objects, as well as the perceptual span of the participant. Given that the distribution of fixations exceed the boundaries of an object, there is a possibility that distributions from different objects overlap. Under conditions of overlapping fixation distributions, we would expect smaller AOI margins to improve model performance as these yield a better ratio of true to false positives. On the other hand, in conditions with no overlap, we would expect better model performance with larger AOI margins, as these minimize the number of false negatives without increasing false positives.

In order to examine our propositions, we reanalyzed four experiments with different AOI margins and conducted a fifth experiment in which we manipulated the distance between choice options. The reanalyses revealed that Studies 1 and 2, which were relatively simple in their visual layout, and where we expected little or no overlap of fixation distributions, benefitted from larger AOI margins. Study 3 used a relatively less accurate eye tracker and large and salient objects with shorter distances between objects possibly leading to overlapping fixation distributions. We found that smaller AOI margins increased the explained variance. Study 4 had high-visual complexity with no distance between objects, and we therefore expected some overlap of fixation distributions. In Study 4, smaller AOI margins (0° of visual angle margin) improved the accuracy of predictions, while the best model fit was obtained with a. 5° of visual angle margin. As a final test of our ideas, we conducted Study 5 manipulating the distance between choice options. We hypothesized that a shorter distance between options would lead to an overlap of fixation distributions, and that this overlap would diminish the gaze bias to the chosen option. As expected, we found that a shorter distance between options reduced the gaze bias, indicating that the distance between objects play a role in determining the degree of overlap of fixation distributions. The overlap of fixation distributions did not, however, influence the evidence accumulation model we fitted, suggesting that the evidence accumulation process, or this particular model, is robust under a putative overlap of fixation distributions. An overview of the five studies and their findings is presented in Table 6.

Additionally to these five studies, we also conducted a survey on the use of AOIs in behavioral eye-tracking research. The survey revealed a lack of consensus regarding the use of AOIs. Researchers indicated using AOIs smaller, the same size, and larger than the object of interest. In addition, most researchers indicated not considering the accuracy of the eye tracker or the perceptual span when defining AOI margins. Perhaps more worrying is that one third of the researchers reported analyzing multiple AOI margins. This does not in itself constitute a problem. However, given that the analysis of multiple AOI margins is rarely reported in published work (Schulte-Mecklenbeck et al., 2015), this could indicate that some proportion of eye-tracking researchers have at some point engaged in data fishing. We note however, that the number of respondents in our survey was relatively small and may not be representative of the average eye-tracking researcher. As such the degree to which this type of behavior is present in eye-tracking research cannot be gleaned from this survey alone.

Implications for behavioral eye-tracking research

Overall, our results support the interpretation of AOI margins as a signal detection problem, and we suggest the following general guidelines according to this interpretation:

- When fixation distributions are expected not to overlap, that is, high accuracy and large distances between objects, it is advisable to keep the AOI margin maximal to include all fixations belonging to the object (reduce false negatives).

DOI: 10.1002/bdm

Table 6. Columns from left: study or condition, hardware producer, eye-tracker accuracy, maximum size of stimuli, minimum distance of objects, size of areas of interest (AOIs), expected overlap of fixation distributions, and findings.

Study	Hardware	Accuracy	Stimuli	Stimuli size	Distance	AOIs	Overlap	Findings
1	Eyegaze	.45°	words	3.64 × 1.81°	4.35°	(0°, .5°, .7°, 1°, 1.5°)	Unlikely	Larger AOI margins increase explained variance.
2.1	Eyegaze	.45°	words	2.13 × 1.11°	2.7°–4.35°	(0°, .5°, 1°, 1.5°, .85 × .21°)	Unlikely	Larger AOI margins increase β's but not explained variance.
2.2			words				Unlikely	
3.1	Tobii 2150	.7°	words	4.8 × 2.5°	2.3°	(0°, .5°, 1°)	Likely	Smaller AOI margins increase explained variance.
3.2			mixed				Likely	
3.3			images	5.6 × 4.8°	2°		Likely	
4	Eyelink II	0.5°	images	31° × 25°,	0°	(0°, .5°, 1°, 1.5°)	Likely	Smaller AOI margins increase explained variance.
5	SMI RED	.4°	images	3.43 × 3.48°	5.89° vs 17.19°	1°	Likely	Larger object distances increase gaze bias effect but not accumulation parameter.

- When fixation distributions are expected to overlap, that is, low accuracy and little distance between objects, it is advisable to use smaller AOI margins (≈0° margin) to balance the ratio of true and false positive fixations.
- When free to design the experimental stimuli, it is advisable to maximize the distance between objects to reduce the overlap of fixation distributions and allow for larger AOI margins that reduce the number of false negative fixations.
- We see two possibilities when deciding on AOI margins. Either the researcher decides a priori based on expectations regarding the overlap of fixation distributions or the AOIs are chosen a posteriori based on how the data complies with one or more quality criteria, such as the achieved measurement accuracy. In either case, eye-tracking studies should report the AOI margin, justify the AOI definition, and report whether more than one AOI was analyzed.

Regarding the practice of analyzing several AOI margins, we must emphasize that this need not be a problem as long as the criterion is maximizing the explained variance and not to produce significance. However, we would urge that the use of multiple AOIs should only be undertaken in pilot studies where plausible assumptions about the distributions of fixations cannot be made with any accuracy. In any case, when several AOI margins are analyzed, details of these analyses should be reported alongside any final analyses.

Limitations and future research

One of the main limitations of our research stems from the fact that we cannot inspect which distribution a particular fixation belongs to. Although our findings are consistent with the interpretation of AOI margins as a signal detection problem we cannot exclude other interpretations. Therefore, an important issue for future research is to devise a method for determining what information is being processed during a given fixation. Given a high fidelity method of determining the information uptake during fixations it would be possible to rule out alternative explanations for our findings.

More generally, there is a distinction to be made between visual processing and the location of fixations. In line with the majority of eye-tracking studies in decision making, and other fields such as visual search and salience, we use AOIs as a method of assigning fixations to objects. The limitation of this approach is the assumption that observers only attend to one object at a time, that visual attention and fixation locations go hand in hand, and that fixations that fall outside AOIs are uninteresting. While these assumptions may hold in some cases, particularly with simple stimuli, they do not hold in general. A common situation in which these assumptions are violated is during visual search when observers fixate to the center of gravity of a subset of search items (Rao et al., 2002). Similar behavior is seen during optimal search models in which saccades are directed to the location that maximizes the probability of being able to correctly identify the target on the next fixation (Najemnik & Geisler, 2008). An area of interest approach would not be suitable for analyzing either model, nor do they allow

DOI: 10.1002/bdm

for attention to be deployed to parafoveal and peripheral regions. In their object-naming study, Clarke et al. (2013) found that a third of named objects had not been fixated. In addition, only half of the fixated objects went on to be named suggesting that not all fixated AOIs are fully processed and encoded to memory.

A second limitation lies in the extensive use of reanalyses. Although this approach provided the opportunity to inspect many more studies, it also limits the inferences that can be drawn as these experiments were not conducted for the purpose of examining the role of AOI margins. In addition, the studies we reanalyzed are by no means inclusive of all the types of decision making research that employ eye tracking. Future studies should ideally conduct experiments with clear data quality criteria so as to derive more specific guidelines for the optimal definition of AOI margins in different visual environments.

One final limitation is that the survey we employed consisted of a small sample of researchers. As such, the data it generated may not accurately capture the diversity in research using eye-tracking methodologies. However, despite the limitations, the survey provides some insight into problems, which may be present in current eye-tracking research. We propose that a more inclusive survey is conducted to provide more certainty about current research practices.

CONCLUSION

In this paper, we examine the influence of AOI definitions on the measurement of object attention and the inferences drawn about cognitive processes. Overall, our results support the interpretation of AOIs as a signal detection problem. We propose a guideline for determining the appropriate AOI definition for a given experiment based on the hardware accuracy and dimensions of the experimental stimuli. Given the increasing use of eye-tracking methods in behavioral research, we encourage future research into how best to use and define AOIs and hope the current work can act as a starting point.

ACKNOWLEDGEMENTS

The authors wish to thank Martin Bagger, Susann Fiedler, Kenneth Holmqvist, Martin Meissner, and Rebecca Wright for lending their time (data).

REFERENCES

Anstis, S.M. (1974). A chart demonstrating variations in acuity with retinal position. *Vision Research*, *14*, 589–592. DOI:10.1016/0042-6989(74)90049-2

Ashby, N. J. S., Dickert, S., & Glöckner, A. (2012). Focusing on what you own: Biased information uptake due to ownership. *Judgment and Decision Making*, *7*(3), 254–267.

Clarke, A. D. F., Coco, M. I., & Keller, F. (2013). The impact of attentional, linguistic, and visual features during object naming. *Frontiers in Psychology*, *4*, 927. DOI:10.3389/fpsyg.2013.00927

Costa-Gomes, M., Crawford, V. P., & Broseta, B. (2001). Cognition and behavior in normal-form games: An experimental study. *Econometrica*, *69*(5), 1193–1235. DOI: 10.1111/1468-0262.00239

Einhäuser, W., Spain, M., & Perona, P. (2008). Objects predict fixations better than early saliency. *Journal of Vision*, *8*(14), 18. DOI: 10.1167/8.14.18

Fiedler, S. & Glöckner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis, *Frontiers in Psychology*, *3*, 335. DOI: 10.3389/fpsyg.2012.00335

Kowler, E. (2011). Eye movements: The past 25 years. *Vision Research*, *51*(13), 1457–1483. DOI: 10.1016/j.visres.2010.12.014

Krajbich, I., Armel, C., & Rangel, A. (2011). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298. DOI: 10.1038/nn.2635

Luce, R. D. (1956). Semiorders and a theory of utility discrimination. *Econometrica, Journal of the Econometric Society*, *178-191*. DOI: 10.2307/1905751

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: Analysis and application*. Cambridge: Cambridge University Press.

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.

Melmoth, D. R., & Rovamo, J. M. (2003). Scaling of letter size and contrast equalises perception across eccentricities and set sizes. *Vision Research*, *43*(7), 769–777. DOI: 10.1016/S0042-6989(02)00685-5

Najemnik, J., & Geisler, W. S. (2008). Eye movement statistics in humans are consistent with an optimal search strategy. *Journal of Vision*, *8*(3), 4. DOI: 10.1167/8.3.4

Nuthmann, A., & Henderson, J. M. (2010). Object-based attentional selection in scene viewing. *Journal of Vision*, *10*(8), 20. DOI: 10.1167/10.8.20.

Orquin, J. L., Bagger, M. P., & Mueller Loose, S. (2013). Learning affects top down and bottom up modulation of eye movements in decision making. *Judgment and Decision Making*, *8*(6), 700–716.

Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, *144*(1), 190–206. DOI: 10.1016/j.actpsy.2013.06.003.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology* (Vol. *1*). Cambridge, MA: MIT press.

Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *Journal of Vision*, *7*(2), 20. DOI: 10.1167/7.2.20

Rao, R. P., Zelinsky, G. J., Hayhoe, M. M., & Ballard, D. H. (2002). Eye movements in iconic visual search. *Vision Research*, *42*(11), 1447–1463. DOI: 10.1016/S0042-6989(02)00040-8

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. DOI: 10.1080/17470210902816461

Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science*, *12*(1), 48–55. DOI: 10.1111/1467-9280.00309

Reingold, E. M., & Sheridan, H. (2011). Eye movements and visual expertise in chess and medicine. *Oxford handbook on eye movements*, *528–550*.

Rogers, W. H. (1993). Regression standard errors in clustered samples. *Stata Technical Bulletin*, *13*, 19–23.

Russo, J. E. (2011). Eye fixations as a process trace. *A handbook of process tracing methods for decision research*, *43–64*.

Schulte-Mecklenbeck, M., Fiedler, S., Renkewitz, F., & Orquin, J.L. (2015) *A review and guideline for the reporting of behavioral*

DOI: 10.1002/bdm

eye-tracking studies. Unpublished dataset. Berlin: Max Planck Institute for Human Development.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, *6*(12), 1317–1322. DOI: 10.1038/nn1150

Simonsohn, U., Nelson, L.D., & Simmons, J.P. (2014) P-curve: A key to the file drawer. *Journal of Experimental Psychology: General*, *143*(2), 534–547. DOI:10.1037/a0033242

Simmons, J.P., Nelson, L.D. & Simonsohn, U. (2011) False-positive psychology: Undisclosed flexibility in data collection and analysis allow presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. DOI: 10.1177/0956797611417632

Steinman, R. M., Haddad, G. M., Skavenski, A. A., & Wyman, D. (1973). Miniature eye movement. *Science*, *181*(4102), 810–819. DOI:10.1126/science.181.4102.810

Svenson, O. (1979). Process descriptions of decision making. *Organizational Behavior and Human Performance*, *23*(1), 86–112. DOI: 10.1016/0030-5073(79)90048-5

Yun, K., Peng, Y., Samaras, D., Zelinsky, G. J., & Berg, T. L. (2013). Exploring the role of gaze behavior and object detection in scene understanding. *Frontiers in Psychology*, *4*, 917. DOI: 10.3389/fpsyg.2013.00917.

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review*, *115*(4), 787–835. DOI: 10.1037/a0013118.

Authors' biographies:

Jacob L. Orquin is an assistant professor at the Department of Business Administration at Aarhus University. His research focuses on eye movements in decision-making and aims to integrate theories from vision and decision research.

Nathaniel J. S. Ashby is a postdoctoral researcher at the Department of Social and Decision Sciences at Carnegie Mellon University. His research focuses on memory processes in decision-making under risk.

Alasdair D. F. Clarke is a postdoctoral researcher at the School of Psychology at the University of Aberdeen. His research is focused on visual search, scene perception, and visual salience.

Authors' addresses:

Jacob L. Orquin, Department of Business Administration/MAPP, Aarhus University, Aarhus, Denmark.

Nathaniel J. S. Ashby, Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA, USA.

Alasdair D. F. Clarke, School of Psychology, University of Aberdeen, Aberdeen, UK.

DOI: 10.1002/bdm